FORM PTO-139 (REV 10-2000)	0 U.S. DEPARTMENT OF COMMERCE PATE	ATTORNEY'S DOCKET NUMBER						
			33339/208804					
	TRANSMITTAL LETTER TO T		U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5)					
 	DESIGNATED/ELECTED OFF CONCERNING A FILING UND	To be assisted / 762194						
INTERNATIONAL APPLICATION NO.		INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED					
PCT/FR99/01908		August 2, 1999	August 4, 1998					
TITLE OF INVENTION								
NUCLEIC SEQUENCES CODING FOR AN AT2 INTERACTING PROTEIN INTERACTING WITH THE AT2 RECEPTOR AND THEIR APPLICATIONS								
APPLICANT(S) FOR DO/EO/US								
ELBAZ, Nathalie; NAHMIAS, Clara; STROSBERG, Arthur, Donny								
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:								
1. 🛛	This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.							
2. 🗌	This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.							
3. 🛛	This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).							
4. 🛛	The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).							
5. 🗵	A copy of the International Application as filed (35 U.S.C. 371(e)(2))							
		(required only if not communicated by the International	al Bureau).					
	 b.							
6. 🗆	A English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).							
7. 🛛	Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))							
	 a. are attached hereto (required only if not communicated by the International Bureau). b. have been communicated by the International Bureau. 							
	<u>—</u>	nicated by the international Bureau. le; however, the time limit for making such amendmen	nts has NOT expired.					
	<u>=</u>	le and will not be made.						
8.	An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).							
9. 🗌	An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).							
10. 🖸	An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).							
Items 11. To 16. Below concern other document(s) or information included:								
11. 🛛	An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98.							
12.	An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.							
13.	A FIRST preliminary amendment. A SECOND or SUBSEQUENT preliminary amendment.							
14. 🔲	A substitute specification.							
15.	A change of power of attorney and/or address letter.							
16. 🗌	Other items or information:							

JC02 Rec'd PCT/PTO 0 5 FEB 2001

U.S. APPLICATION NOTIFICATION See 7 CGR 201 9 4 INTERNATIONAL APPLICATION NO. PCT/FR99/01908			ATION NO.	ATTORNEY'S DOCKET NUMBER 33339/				
17. The following fees are submitted:	CALCULATIONS	PTO USE ONLY						
Basic National Fee (37 CFR 1.492(a)(1)-(5)) Neither international preliminary examination search fee (37 CFR 1.445(a)(2)) paid to USPTo not prepared by the EPO or JPO								
International preliminary examination fee (37 of Search Report prepared by the EPO or JPO International preliminary examination fee (37 search (37CFR 1.445(a)(2)) paid to USPTO International preliminary examination fee (37 of But all claims did not satisfy provisions of PC International preliminary examination fee (37 of provisions of PCT Article 33(1)-(4)								
ENTER APPROPRIATE BASIC FI			= 20 30 months	\$ 860.00				
Surcharge of \$130.00 for furnishing the or from the earliest claimed priority date (37)	\$	·						
CLAIMS NUMBER FIL		NUMBER EXTRA	RATE	\$ 0.00				
Total Claims 20 -20 Independent Claims 2 - 3		0	X \$18.00 X \$80.00	\$ 0.00				
Independent Claims 2 - 3 MULTIPLE DEPENDENT CLAIM(S) (if appli			+ \$270.00	S				
		L OF ABOVE CALO	CULATIONS =	\$ 860.00				
Applicant claims small entity status.								
reduced by one-half.				\$				
			SUBTOTAL =	\$ 860.00				
Processing fee of \$130.00 for furnishing the months from the earliest claimed priority of	\$							
		TOTAL NA	TIONAL FEE =	\$ 860.00				
Fee for recording the enclosed assignment (37 by an appropriate cover sheet (37 CFR 3.28, 3.28).	\$							
	\$ 860.00							
	Amount to be Refunded	s						
				Charged	S			
a A check in the amount of	Charged							
 a. A check in the amount of \$ 860.00 to cover the above fees is enclosed. b. Please charge my Deposit Account No. 16-0605 in the amount of \$ to cover the above fees. A duplicate copy of this sheet is enclosed. 								
c. The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 16-0605.								
Note: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.								
SEND ALL CORRESPONDENCE TO: Raymond O. Linker, Jr.		"Express Mail" Mail Date of Deposit: Febr	"Express Mail" Mailing Label Number EL 432823389 US Date of Deposit: February 5, 2001					
SIGNATUKE REGISTRATION NUMBER 26,419 ALSTON & BIRD LLP Post Office Drawer 34009 Charlotte, NC 28234)	Postal Service "Expres	Thereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to BOX PCT, Attn: DO/US (PTO) Commissioner for Patents, Washington, DC 20231.					
Tel. Charlotte Office (704) 331-6000 Fax Charlotte Office (704) 334-2014 Customer Number 000826	+ JUNIO PO							

CLT01/4463187v1

Attorney's Docket No. 33339/208804

PATENT

IN THE UNITED STATES DESIGNATED OFFICE (DO/US)

In re:

Attn: DO/US

International Appl. No.

PCT/FR99/01908

International Filing Date:

August 2, 1999

For:

NUCLEIC SEQUENCES CODING FOR AN AT2 INTERACTING PROTEIN INTERACTING

WITH THE AT2 RECEPTOR AND THEIR APPLICATIONS

STATEMENT IN SUPPORT OF FILING A SEQUENCE LISTING UNDER 37 CFR § 1.821(f)

Box PCT Commissioner for Patents Washington, DC 20231

Sir:

I hereby state that the content of the paper and computer readable copies of the Sequence Listing, submitted concurrently herewith in accordance with 37 CFR § 1.821(c) and (e), are the same.

Respectfully submitted,

Raymond O. Linker, Jr. Attorney/Agent for Applicant Registration No. 26,419

Alston & Bird LLP

Bank of America Plaza 101 South Tryon Street, Suite 4000 Charlotte, NC 28280-4000 Tel Charlotte Office (704) 444-1000 Fax Charlotte Office (704) 444-1111

"Express Mail" Mailing Label Number EL 836091314 US Date of Deposit: April 19, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Box PCT, Commissioner for Patents, Washington, DC 20231.

Sarah B. Simmons

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner For Patents, Washington, DC 20231, on April 19, 2001.

WO 00/08148

10

25

30

35

NUCLEIC SEQUENCES ENCODING AN AT2 RECEPTOR-INTERACTING PROTEIN (ATIP) AND THEIR APPLICATIONS

The present invention relates to nucleic sequences encoding a protein capable of interacting with the AT2 receptor, to oligonucleotides contained in the said sequences, to their applications as probes and for the expression of the said proteins, to the vectors useful for the said expression, to the host cells containing the said vectors and to a model for studying the AT2 receptor.

The present invention also relates to the said proteins and to their applications. $\dot{}$

The octapeptide, angiotensin II, mainly known as a regulator of blood pressure, has also been described as an important modulator of cell growth. Interestingly, this peptide appears to exert opposite effects on cell growth according to whether it is bound to one or the other of its two subtypes of membrane receptors (AT1 or AT2).

The AT2 receptor subtype, which also belongs to the G protein-coupled receptor family, is still poorly characterized both from the point of view of its mechanisms of activation and its physiological role (C. Nahmias et al., Trends Pharmacol Sci, 1995, 16, 223-225). Several arguments suggest, however, a role for this receptor in the phenomena of cell proliferation, differentiation or adhesion.

The AT2 receptor is highly expressed during foetal life, disappears in adults in most tissues, but becomes reexpressed under pathophysiological conditions involving restructuring of the tissues.

Studies carried out *in vivo* have demonstrated the inhibitory role exerted by the AT2 subtype on the proliferation of the muscle cells of the *tunica intima vasorum* after vascular lesion (P. Janiak et al., *Hypertension*, 1992, 20, 737-745; M Nakajima et al., *Proc. Natl. Acad. Sci.* USA, 1995, 92, 10663-10667).

Moreover, the stimulation of the AT2 receptor activates phosphatase SHP-1 (Bedecs K., et al; Biochem. J., 1997, 325, 449-454). The fact that the AT2 receptor activates a phosphatase is consistent with its antiproliferative effects.

In the light of the above, it has been shown that, on cells in culture, the AT2 receptor:

- inhibits the synthesis of DNA and proliferation, which are induced by angiotensin II (Ang II) and bFGF (M. Stoll et al., *J. Clin. Invest.*, 1995, 95, 651-657),

10

20

- induces apoptosis (T. Yamada et al., Proc. Natl. Acad. Sci. USA, 1996, 93, 156-160), and
- induces neuronal differentiation (L. Laflamme 15 et al., J. Biol. Chem., 1996, 271, 22729-22735).

Studies of the signalling pathways associated with the AT2 receptor have been undertaken in cells of the N1E-115 line which are derived from a murine neuroblastoma and which express only the AT2 subtype. A

- first study has made it possible to demonstrate rapid and transient dephosphorylation of some proteins on the tyrosine residues following the treatment of N1E-115 cells with angiotensin II (C. Nahmias et al., Biochem. J., 1995, 306, 87-92). It has also been shown that the
- 25 AT2 receptor interferes with the pathways for activation of growth factor receptors and inhibits the activity of MAP kinases (ERK1 and ERK2) (mitogenactivated protein), which play a key role in the phenomena of cell proliferation and differentiation.
- The inhibitory effect of AT2 on the activation of MAP kinases is rapid and transient, does not involve a regulatory protein sensitive to the pertussis toxin (of the Gi/Go type), but involves the activation of an orthovanadate-sensitive tyrosine phosphatase.
- Taking into account the role of the AT2 receptor in cell proliferation, the inventors have sought to develop tools capable of regulating the action of the AT2 receptor. Indeed, the activation of

the AT2 receptor may have repercussions in cancerology (inhibition of cell proliferation).

In general, the AT2 receptor has opposite effects to those of AT1 on the activation of MAP kinases and on cell proliferation; study of the communication which may exist between these two receptor subtypes, which bind the same ligand, is consequently of interest.

5

10

15

20

25

30

35

The study of the signalling pathways and of the regulation of the AT2 receptor also represents a major stake for human health knowing that antagonists of the AT1 receptor are currently administered to patients with hypertension. In this context, it is essential to know the biological effects associated with the AT2 receptor which remains activable by circulating Ang II in this type of treatment.

The subject of the present invention is an isolated nucleic acid (DNA or RNA) fragment, encoding a protein capable of binding to the AT2 receptor, which fragment is selected from the group consisting of the sequences SEQ ID NO:1, 3, 5, 7 and 9, as represented in the sequence listing included in the present application.

These various sequences correspond to the complementary DNA (cDNA) encoding all or part of the protein called hereinafter ATIP (AT2 interacting protein).

The sequence SEQ ID NO:1 (1803 bp) corresponds to the complete nucleic sequence of mouse ATIP and includes both the parts encoding the AT2 receptor binding protein and the noncoding parts.

The sequence NO:3 (1323 bp) corresponds to the nucleic acid sequence of the coding part of the sequence SEQ ID NO:1, while the sequence SEQ ID NO:5 corresponds to the sequence NO:1 fragment obtained by the two-hybrid technique (A Plessis et al., M/S, 1994, 9, I-1K; J. Luban et al., Curr. Op. Biotechnol., 1995, 6, 59-64).

The sequence SEQ ID NO:7 (3742 bp) corresponds to the complete nucleic sequence of the human cDNA and includes both the parts encoding the protein homologous to the mouse ATIP and the noncoding parts.

The sequence SEQ ID NO:9 (1308 bp) corresponds to the coding part of the sequence SEQ ID NO:7.

5

10

15

20

25

30

35

The subject of the present invention is also transcripts, characterized in that they are complementary to the sequences in accordance with the invention and are in particular generated from the said sequences.

The subject of the present invention is, in addition, fragments of the said sequences comprising between 20 and 400 bp, useful as probes or as primers, for the detection of the sequences SEQ ID NO:1, 3, 5, 7 or 9, or of homologous sequences.

Among the said fragments, there may be mentioned in particular a probe of 354 bp (SEQ ID NO:5) as well as any fragment of 20 bp to 400 bp included in the sequences SEQ ID NO:1, 3, 5, 7 or 9.

As primer, there will be used in particular the sequence SEQ ID NO:10 (antisense oligonucleotide) which makes it possible in particular to amplify the 5' parts of the various mRNAs corresponding to ATIP (5' RACE technique: Marathon cDNA amplification kit, Clontech).

It is also possible to use, as amplification primers, any pair of oligonucleotides of more than 20 bp and comprising part of the ATIP (human or mouse) nucleic sequence, in particular the pair SEQ ID NO:11-SEO ID NO:12.

The preferred hybridization (prehybridization and hybridization) conditions are in particular the following: 45% formamide, 9% dextran sulphate, 0.2% BSA, 0.2% polyvinyl pyrrolidone, 0.2% Ficoll, 0.1% sodium pyrophosphate, 0.01% SDS, 0.05 mM Tris pH 7.5, 0.9 M NaCl and rinses to a stringency corresponding to the buffer: 1XSSC, 0.1% SDS.

The subject of the present invention is also a purified and isolated protein, called ATIP, which is capable of interacting with the AT2 receptor and which is selected from the group consisting of the sequences SEQ ID NO:2, 4, 6 or 8.

5

10

15

20

25

30

35

The murine and human sequences exhibit 85.6% homologies. The human sequence (human ATIP) possesses 5 amino acids less than the mouse sequence (mouse ATIP). The amino acids missing from the human sequence are situated at the level of amino acids: 162, 163, 164, 166 and 214 of the mouse ATIP sequence.

Comparisons (Blast) between the ATIP protein sequences according to the invention and the sequences contained in data banks indicate that human ATIP (like mouse ATIP) never exhibits more than 25% homology with a known sequence, and this being the case only over part of this sequence.

The subject of the present invention is also a translational product, characterized in that it is encoded by a nucleotide sequence in accordance with the invention.

The subject of the present invention is, in addition, antibodies, characterized in that they are directed against the ATIP protein or an ATIP protein fragment according to the invention.

The subject of the present invention is also a recombinant cloning and/or expression vector, characterized in that it comprises a nucleotide sequence in accordance with the invention.

The subject of the present invention is also a transformed host cell, characterized in that it comprises a vector as defined above.

Among the preferred transformed cells according to the invention, there may be mentioned $E.\ coli$ and CHO cells.

The subject of the present invention is also transformed host cells, characterized in that they consist of a suitable yeast strain cotransformed with

at least two vectors which respectively encode (i) a so-called bait protein selected from the consisting of a fragment containing at least SEO ID NO:5 of the ATIP protein and a fragment containing at least the C-terminal end of the AT2 receptor, which bait protein is fused with a protein selected from the group consisting of the DNA-binding domain of transcription factor and the activation domain of the same transcription factor and (ii) a so-called prey protein, selected from the group consisting of a fragment containing at least SEQ ID NO:5 of the ATIP protein, a fragment containing at least the C-terminal end of the AT2 receptor and any other polypeptide corresponding to a sequence contained in a cDNA library, which prey protein is fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the same transcription factor, which vectors comprise, in addition, selectable markers.

10

15

20

25

30

35

According to an advantageous embodiment of the said cells, they consist in particular of:

- either a suitable yeast strain cotransformed with three vectors which respectively encode (i) a bait corresponding to a fragment containing the C-terminal end of the AT2 receptor fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, (ii) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to the invention, fused with a protein selected from the group consisting of the DNA-binding domain transcription factor and the activation domain of the transcription factor and (iii) a polypeptide corresponding to a sequence contained in library, which vectors comprise, addition, in selectable markers,
- or a suitable yeast strain cotransformed with two vectors which respectively encode (i) a fragment

containing at least SEQ ID NO:5 of the ATIP protein according to the invention, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a polypeptide corresponding to a sequence contained in a cDNA library, fused with a protein selected from the group consisting of the DNA-binding domain of the transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers,

10

15

20

25

30

35

- or a suitable yeast strain cotransformed with two vectors, namely (i) a vector encoding a fragment containing at least SEQ ID NO:5 of the ATIP protein, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a vector encoding a fragment containing the C-terminal end of the AT2 receptor, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, one of the two vectors necessarily encoding a mutated protein.

The subject of the present invention is also a method for selecting proteins inhibiting ATIP protein according to the invention-AT2 receptor interaction, which method comprises:

(a) cotransforming a suitable yeast strain with three vectors which respectively encode (i) a bait corresponding to a fragment containing the C-terminal end of the AT2 receptor fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, (ii) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to the invention, fused with a protein selected from the

group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (iii) a polypeptide corresponding to a sequence contained in a cDNA library, which vectors comprise, in addition, selectable markers,

5

10

15

20

25

- (b) selecting the clones of cDNA library expressing a polypeptide inhibiting the AT2 receptor-ATIP protein according to the invention interaction, on an appropriate selective medium, and
 - (c) identifying the said polypeptide.

Such a method uses in particular the so-called reverse two-hybrid or three-hybrid technique as described in Vidal et al. (*Proc. Natl. Acad. Sci.* USA, 1996, 93, 10315-10320 and 10321-10326) or Tirode et al. (*J. Biol. Chem.*, 1997, 272, 37, 22995-22999).

The subject of the present invention is also a method for screening polypeptides interacting with the ATIP protein according to the invention, which method comprises:

- (a) cotransforming a suitable yeast strain with two vectors as defined above, namely which respectively encode (i) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to the invention, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a polypeptide corresponding to a sequence contained in a cDNA library, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, and
- (b) selecting the clones expressing a 35 polypeptide interacting with the ATIP protein, on a suitable selective medium.

Such a method makes it possible in particular to search for other proteins interacting with the ATIP

protein, in particular in order to find the next links in the pathway activated by the AT2 receptor, so as to use them to modify the protein according to the invention-AT2 receptor interaction.

The subject of the present invention is also a method for characterizing the domains involved in the ATIP protein-AT2 receptor interaction, characterized in that it comprises:

5

10

15

20

30

35

- (a) cotransforming a suitable yeast strain with two vectors, as defined above, namely (i) a vector encoding a fragment containing at least SEQ ID NO:5 of the ATIP protein, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a vector encoding a fragment containing the C-terminal end of the AT2 receptor, mutated or not, fused with a protein selected from the group consisting of the DNAbinding domain of a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, one of the two vectors necessarily encoding a mutated protein, and
- (b) visualizing, by selection on a suitable 25 selective medium, the possible loss of the ATIP-AT2 receptor interaction.

Such a method makes it possible to identify and to delimit the important domains of the ATIP protein or of the C-terminal end of the AT2 receptor, on which their interaction depends, so as to use them as preferred target for modifying the AT2 receptor signalling.

The subject of the present invention is also a method for selecting substances capable of influencing the ATIP protein according to the invention-AT2 receptor interaction, which method comprises:

(a) bringing the ATIP protein, attached to a support, into contact with a fusion protein AT2

receptor-protein tag, optionally in the presence of a substance to be tested,

(b) at least one washing of the said support thus treated with a suitable buffer, and

5

15

20

25

30

35

(c) visualizing the possible ATIP-AT2 receptor interaction, in particular in SDS-PAGE, followed by immunoblotting with antibodies directed against the protein tag, fused with the AT2 receptor, or against the AT2 receptor.

If the substance to be tested inhibits the ATIP-AT2 receptor interaction, the visualization step is negative.

In accordance with the invention, ATIP is attached to the said support either covalently, or through affinity binding between an attachment substance fused with ATIP and the said support. For example, the said support consists of beads coupled either to a substance having affinity with the said attachment protein, fused with ATIP, or to suitable antibodies.

The fusion protein AT2 receptor-protein tag is in particular obtained from a lysate of cells transfected with a vector expressing the fusion protein AT2-protein tag.

As a variant, the said method for selecting substances capable of interacting with the ATIP protein according to the invention comprises:

- (a) bringing the ATIP protein, attached to a support, into contact with a cell lysate,
- (b) at least one washing of the said support thus treated with a suitable buffer,
- (c) visualizing the possible protein combined with the ATIP protein, in particular in SDS-PAGE, followed by immunoblotting with appropriate antibodies, and
- (d) identifying the protein in the cell lysate interacting with the ATIP protein.

In accordance with the said method for selecting substances capable of influencing the ATIP protein according to the invention-AT2 interaction, it is possible to use in particular, as fusion proteins ATIP-protein tag, the proteins GST-ATIPc and MYC-ATIPc, which constitute tools which can make it possible to bring down in vitro any proteins interacting with ATIP, for example, from cell lysates activated or otherwise with ligands for the AT2 receptor. The GST-ATIP protein may be brought down by specific interaction of GST with agarose beads coupled glutathione, or alternatively immunoprecipitated with the anti-ATIP antibody. The Myc-ATIP protein may immunoprecipitated with commercial anti-MYC antibodies or with the anti-ATIP antibody.

10

15

20

25

30

The advantage of these methods consists in finding means of modifying the signalling, the level of expression or the pharmacology of the AT2 receptor, which may have therapeutic applications. Indeed, when a pathological condition has been clearly correlated with a transduction abnormality associated with the AT2 receptor, modification of this transduction, in particular by acting on the binding of the AT2 receptor to the protein according to the invention, may then possibly compensate for the pathological disorder or at least influence it.

The subject of the present invention is also the use of the abovementioned cotransformed cells for the selection and screening of substances or of proteins capable of influencing the ATIP protein-AT2 receptor interaction or capable of interacting with the ATIP protein.

In addition to the preceding features, the invention also comprises other features which will emerge from the description which follows, which refers to exemplary embodiments of the method which is the subject of the present invention as well as to the accompanying drawings, in which:

- Figure 1 corresponds to the C-terminal end of the mouse AT2 receptor, used as a two-hybrid bait for screening a mouse cDNA library;
- Figure 2 illustrates the position of the
 GAL4-binding domain and the multiple cloning site of the plasmid pGBT9 (Clontech);
 - Figure 3 illustrates the presumed coiled-coil structures (coiled-coil domains underlined) of mouse ATIP;
- Figure 4 illustrates the presumed coiled-coil structures (coiled-coil domains underlined) of human ATIP;
 - Figure 5 illustrates the structure of the plasmid pVP16;
- Figure 6 illustrates the multiple cloning site of the plasmid pRSET A;
 - Figure 7 illustrates the MCY sequence used to construct the plasmid pcDNA3-MYC;
- Figure 8 illustrates the structure of the 20 plasmid pBAC-PAK-poly HIS;
 - Figure 9 illustrates a Northern blot of several human tissues hybridized with the probe ATIPmouse-short (SEQ ID NO:5);
- Figure 10 illustrates the interaction in vitro of the protein ATIPmouse-short with the C-terminal end of the AT2 receptor; and
 - Figure 11 illustrates the modifications of the signal induced by the AT2 receptor by overexpression of the ATIP protein.
- It should be clearly understood, however, that these examples are given solely by way of illustration of the subject of the invention and do not constitute in any manner a limitation thereto.

EXAMPLE 1: Demonstration of a specific protein-protein interaction between the AT2 receptor and the protein having the sequence SEQ ID NO:6 according to the invention

Materials and methods

5

10

15

30

35

- The two-hybrid system, initially developed by Song and Fields in 1989 (Nature, **340**, 245-246) is based on the fact that the activity of numerous eukaryotic transcription-activating factors requires only two domains: an activating domain which does not have the capacity to bind DNA and a DNA-binding domain.

In the two-hybrid system, the DNA-binding domain is fused with a protein X and the activation domain is fused with a protein Y. If, and only if, X and Y interact, a complex is formed which reconstitutes a functional transcription factor.

- Construction of the expression vectors:
- . "bait" vectors:

Protein X: C-terminal end of the sequence encoding the mouse AT2 receptor (52 amino acids of CVNPF at the stop codon, see Figure 1), fused with the sequence encoding the Gal4 DNA-binding domain (Figure 2).

Insert: end of the mouse AT2 receptor (159 bp + 25 16 bp of sites generated by PCR) inserted at the level of the EcoRI and BamHI sites of the vectors pLEX9 (Clontech) or pGBT9 (modified pGAD424 or pBTM116; A.B Vojtek et al., Cell, 1993, 74, 205-214).

The following sequence is thus obtained: CGGAATTC on the 5' side-AT2 C-terminal sequence of 52

. screened library:

amino acids-GGATCCCG 3' side

mouse foetal cDNA library (A.B. Vojtek et al., Cell, 1993, 74, 205-214), containing inserts of 350 to 700 bp (protein Y) in the vector VP16 (Figure 5).

. "Bait" control vectors

Protein X: C-terminal end of the human $\beta 2$ -adrenergic receptors, rat AT1 or human bradykinin.

. Transformed yeast strain

HF7c (Clontech) for the bait constructed in pGBT9;

L40 for the bait constructed in pLex9.

Results

5

10

15

20

25

30

35

This strategy made it possible to isolate a clone derived from the cDNA library containing an insert of 354 bp (ATIP) which interacts specifically with the C-terminal end of AT2. It is of interest to note that the screening of this library with the constructs produced in the two expression vectors pGBT9 and pLEX9 made it possible to find this same clone in both cases. This clone does not interact with control proteins exhibiting nonspecific interactions.

the selectivity of this To evaluate interaction, the ATIP clone was tested as a two-hybrid system with the C-terminal ends of the receptors: human β 2 adrenergic, rat AT1 and human bradykinin, and all negative results. This indicates that polypeptide encoded by the ATIP clone interacts, in a selective manner, with the C-terminal end of the mouse AT2 receptor.

EXAMPLE 2: Characterization of the ATIP clone

To test for the corresponding whole clone, a probe of 354 bp (SEQ ID NO:5), which corresponds to the insert obtained by digestion with the restriction enzyme NotI of the plasmid isolated in a two-hybrid system (that extracted from the VP16 library, selected as being positive in the screen using, as bait, the Cterminal end of the mouse AT2 receptor), is used to screen a mouse foetal cDNA library constructed with inserts of more then 1 kb in size. Two overlapping comprising the ATIP sequence, were identified and made it possible to sequence 1803 bp of the corresponding cDNA (SEQ ID NO:1). This sequence contains an open reading frame of 1323 bp (SEQ ID NO:3), potentially encoding a protein of 440 amino acids (SEQ ID NO:2 and 4). Comparisons between the

identified protein sequence and the sequences contained in data banks indicate that it never exhibits more than 25% homology with a known sequence part.

The 354 bp probe (SEQ ID NO:5) was used as probe in Southern and Northern in a very satisfactory manner under the hybridization conditions below: prehybridization and hybridization in 45% formamide, 9% dextran sulphate, 0.2% BSA, 0.2% polyvinylpyrrolidone, 0.2% Ficoll, 0.1% sodium pyrophosphate, 0.01% SDS, 0.05 mM Tris pH 7.5, 0.9 M NaCl and rinses to stringency: 1 x SSC, 0.1% SDS.

5

10

15

20

25

30

35

In parallel, Northern blot hybridization experiments carried out on total RNAs of N1E-115 cells with the ATIP probe (SEQ ID NO:5) confirm the expression of the corresponding mRNA in the N1E-115 cells, and indicates the existence of at least 5 transcripts of different sizes. These transcripts correspond to alternative splicings of the same gene or to different homologous genes.

On a Northern, performed under the conditions described in the literature on a 5 μ g sample of poly A+RNA of N1E-115 cells, the sizes of the various transcripts hybridizing with the ATIPmouse probe are = 2.5-3.5-5-5.3 and 7.5 kb.

Figure 9 represents a Northern blot containing poly A+ RNAs of various human tissues, hybridized with the same ATIPmouse probe. It is possible to observe that ATIP is ubiquitously expressed. A predominant transcript at 4.4 kb is found in all the tissues represented, to which there are added, according to the tissues, other longer transcripts (pancreas and heart) or shorter transcripts (pancreas, skeletal muscle, placenta, brain and heart). These are perhaps the fruit of an alternative splicing of the ATIP RNA which would be dependent on the tissue considered or alternatively they are the sign of the existence of an RNA family encoding proteins of the "ATIP family" homologous to

ATIP and which are revealed by the probe, at the stringency used.

To know the size of the smallest transcript encoding ATIP, a rapid amplification of the cDNA ends (5' RACE, Marathon cDNA Amplification Kit from Clontech) from poly A+ RNA of N1E-115 cells was carried out using the antisense oligonucleotide of SEQ ID NO:10, to amplify the 5' parts of the various mRNAs corresponding to the endogenous ATIP of the N1E-115 cells (murine neuroblastoma).

5

10

The results obtained indicated that the smallest transcript including the ATIP domain is an mRNA of 1950 bp, which indeed contains the start of the coding sequence obtained by cloning.

Any other pair of oligonucleotides (primers) of more than 20 bp and comprising part of the ATIP sequence may also be used to amplify, by PCR (PCR conditions to be determined for each pair of oligonucleotides with the aid of the OLIGO 4 software), part of the ATIP (and to give a DNA fragment which may

be optionally used as a probe to recognize the DNA or the RNA corresponding to the ATIP).

EXAMPLE 3 Construction of various vectors according to the invention

In general, the vectors containing ATIPmouseshort (with the exception of pRSETA-ATIPmouse-short) were obtained from an insert produced by PCR with the following two oligonucleotides (SEQ ID NO:11 and SEQ ID NO:12):

oligo. sense: 5' CGCGGATCCCAGACAGACCGGACGGAACTGGAG3' oligo. antisense: 5'CCGGAATTCACTACAACCTTTCGTTTAAAGCATC 3',

using as template the vector VP16-ATIPmouseshort (Figure 5). For the sake of convenience, this vector is called ^BATIPc^{stop,E}. Indeed, digested with BamHI and EcoRI, it gives an insert corresponding to the sequence 1st strand: GATCC-SEQ ID NO:5 (minus CAT)-TAGTG
2nd strand: CCTAG------CTTAAG

(STOP)_

BamHI site

5

10

15

20

EcoRI site

Other vectors may also be constructed; they comprise all or part of the ATIP protein and are the following:

-VP16-ATIPmouse-short (vector taken from the library screened in the two-hybrid system, comprises 354 bp (SEQ ID NO:5), inserted in NotI into VP16).

-pCDNA3-MYC-ATIPmouse-short (insert BATIPcstop,E, inserted in BamHI-EcoRI into pCDNA3-MYC (pcDNA3 from Invitrogen, modified by insertion of the MYC sequence, Figure 7); this plasmid may be used in stable or transient transfections. It makes it possible to express MYC-ATIPmouse-short in eukaryotic cells. The expression of this protein in eukaryotic cells after transfection of the corresponding plasmid has already been obtained and checked by immunoreaction with an anti-MYC and anti-ATIP antibody.

-pRSETA-HIS-ATIPmouse-short (insert BATIPcstop,E, inserted in BamHI-EcoRI into pRSETA, Invitrogen). This plasmid makes it possible to express the fusion protein HIS-ATIPmouse-short in bacterial cells and to purify it on a nickel column (see Figure 6 for the multiple cloning site).

-pBacPAK-polyHIS-ATIPmouse-short (insert BATIPcstop, inserted in BamHI-EcoRI into the vector 25 (commercial pBacPAK, modified pBacPAK-polyHIS insertion of a sequence containing a histidine tag and a site for cleavage with thrombin, Figure 8). construct may be used to express the ATIPmouse-short protein, fused with a histidine tag, in insect cells 30 (SF9 type). Indeed, as indicated, this vector contains a poly-histidine insert and can therefore encode the fusion protein. The latter, like the fusion protein cloned into pRSET, may be purified on a nickel column and may serve in the same type of techniques. 35

-pGEX-4T1-GST-ATIPmouse-short (insert amplified by the PCR identical to BATIPcstop, but with no STOP codon, which extends the ATIPmouse-short sequence by the few amino acids which follow: Phe-Glu-Phe-Pro-Gly-Arg-Leu-Glu-Arg-Pro-His-Arg-Asp obtained from the plasmid pGEX-4T-1 (Pharmacia). This plasmid makes it possible to express the protein GST-ATIPmouse-short in bacterial cells and to purify it on glutathione-agarose beads.

-pCDNAI-ATIPmouse clonel (entire 5' sequenced from ATIP and ORF up to bp: 1205 starting from the beginning of the clone, inserted in BstxI into pCDNAI). This plasmid is derived from the cloning of the mouse foetal library with the probe SEQ ID NO:5. This plasmid can serve to produce, in bacteria, the 5' portion of the ATIPmouse DNA, so as to use it as a probe.

10

15

20

25

30

35

-pCDNAI-ATIPmouse clone2 (2nd half of the ORF of ATIP from bp: 616 and up to the end of the 3'sequenced (bp 1803), inserted in BstxI into pcDNAI, Invitrogen). This plasmid can serve to produce, in bacteria, the 3' portion of the ATIPmouse DNA, so as to use it as a probe.

-pCDNAI-ATIPmouse-long (clones 1 and 2 placed end to end, using the intermediate SapI site. This plasmid contains the entire ATIPmouse clone, inserted in BstxI into pCDNAI). This plasmid may be used in transient transfections in eukaryotic cells.

-pcDNA3-ATIPmouse-long (whole ATIPmouse from BamHI-XbaI of pcDNAI-ATIPmouse-long, and inserted into pcDNA3, Invitrogen, at these same sites). This plasmid may be used in stable or transient transfections in eukaryotic cells. It made it possible to translate in vitro (kit TNT T7 coupled reticulocytes lysate systems, Promega) the whole ATIP protein and to observe that its translational product has an apparent molecular weight on gel of 58 kDa. Added to this predominant product are two minor products of 30 and 15 kDa. According to the ATIP sequence, these could correspond to partial

products of translation in vitro starting with ATGs other than that at position 178 of SEQ ID NO:1.

EXAMPLE 4: Production of stable clones expressing the ATIPmouse-short or long protein

Stable clones expressing both the human AT2 receptor and ATIPmouse-short (SEQ ID NO:6) or ATIPmouse-long (SEQ ID NO:3) were obtained by transfection.

5

15

20

25

30

CHO cells, deficient in dihydrofolate reductase, are transfected with a plasmid containing the region encoding the human AT2 receptor (Bedecs et al., Biochem. J. 1997, 325, 449-454).

The clone selected, CHO-hAT2, expressing 100 fmol of AT2 receptor/mg of protein, is cultured on an HAMF12 medium supplemented with 10% foetal calf serum and used between passages 10 and 30.

This clone was itself transfected with the plasmids pCDNA3-MYC-ATIPmouse-short or pCDNA3-ATIPmouse-long described in Example 3. The selection of the clones stably expressing the ATIP protein (short form or long form) was carried out in a selective medium containing 800 μ g/ml of G418. The cell lysates, corresponding to the various selected clones, were subjected to SDS-PAGE followed by immunoblotting and this was incubated with the anti-ATIP polyclonal antibody. The results obtained indicate that various clones expressing various levels of ATIPmouse-short were able to be obtained.

EXAMPLE 5: Production of polyclonal antibodies directed against the SEQ ID NO:6 sequence

To progress in the characterization of this clone, the production of polyclonal antibodies directed against the ATIP domain was undertaken.

For that, a vector encoding a protein corresponding to this domain fused with six histidine residues was constructed.

The following sequence:
GGA TCC-SEQ NO:5-TAG-TGA-ATT

is inserted into the plasmid pRSETA, as defined above.

In this insert, SEQ ID NO:5 does not comprise the first CAT.

The plasmid obtained is expressed in the E. coli strain BL 21 (DE3) (F ompT r_B m_B) containing the bacteriophage DE3 which carries a DNA fragment containing the lacI gene, the lacUV5 promoter, the start of the lacZ gene and the gene encoding T7 RNA polymerase. This fragment is introduced into the int gene.

5

10

20

25

30

35

In the presence of DE3, only the lacUV5 promoter, inducible by IPTG directs the transcription of T7 RNA polymerase.

The addition of 0.4 mM IPTG to a culture of BL21 (DE3) cells induces the production of T7 RNA polymerase which, in turn, causes the transcription of the target DNA of the plasmid pRSETA (which allows the translation of the protein binding to the AT2 receptor).

The protein obtained (17 kDa) is purified on a nickel column (Ni-NTA, QuiAexpressionist 07/97, Quiagen), by virtue of the affinity of its six histidine residues for nickel. The protein obtained is then injected into rabbits so as to obtain polyclonal antibodies directed against the ATIP protein. The bleedings obtained have a very good titre.

These antibodies, purified on a GST-ATIP column, after passing through a GST column alone (so as to remove possible GST-specific antibodies and to retain on the GST-ATIP column only the antibodies specific for ATIPmouse-short) may be used successfully to immunoprecipitate and reveal in immunoblotting MYC-ATIPmouse-short from transiently transfected COS cells. Furthermore, this purified antibody also reveals in immunoblotting the ATIPmouse-long protein contained in lysates of COS cells transiently transfected with the plasmid pCDNA3-ATIPmouse-long.

The transfected ATIPmouse-long protein is visualized after SDS-PAGE and immunoblotting with an anti-ATIP antibody, in the form of two polypeptides having apparent molecular weights of 50 and 45 kDa.

5

10

15

20

25

30

35

antibody was purified This immunofluorescence on CHO-hAT2 cells, fixed by a 15minute treatment with paraformaldehyde (3%). After successively treated with the cells are fixing, for solutions of PBS/glycine 50 mΜ 20 minutes, PBS/Triton X100 0.1% for 5 minutes and PBS/BSA 0.2% for 15 minutes. They are then successively incubated in solutions containing 15 µg/ml of antibody containing the purified anti-ATIP antibody, and then the antirabbit immunoglobulin antibody coupled to rhodamine for 30 minutes. Between each new incubation, three rinses under Observations in PBS are carried out. fluorescence microscope indicate an expression of the endogenous ATIP protein in the nucleus (predominantly) and in the cytoplasm of the CHO-hAT2 cells.

Some cells show a homogeneous distribution of the fluorescence due to the anti-ATIP antibody in these compartments, whereas other cells which appear more spread out, show a heterogeneous distribution of the fluorescence along the filaments which appear to start from the nucleus and spread up to the plasma membrane of the cell, in an organized network. Additional colocalization experiments should be carried out to determine if these filaments coincide or otherwise with known structures of the cytoskeleton.

EXAMPLE 6: Confirmation of the in vitro interaction of the ATIPmouse-short protein with the C-terminal end of the AT2 receptor

To demonstrate the interaction of the ATIPmouse-short protein with the C-terminal end of the AT2 receptor in a system other than that of the two-hybrid system, a protocol which makes it possible to demonstrate this interaction in vitro was set up. For that, the fusion protein GST-ATIP as described above

was produced; it is combined through its GST part with glutathione coupled to agarose beads (GA). In parallel, bacteria (DH5α) are transfected with a plasmid (pMALc2-AT2, derived from pMAL-c2 from New England Biolabs) encoding a fusion protein between the C-terminal end of 5 the human AT2 receptor (Asn314-Ser363) and MBP (Maltose Binding Protein). These bacteria were cultured and the fusion protein was induced in 0.3 mM IPTG according to the protocol "pMAL Protein Fusion and Purification System" from New England Biolabs. After centrifugation 10 of the culture at 4 000 g and solubilization of the pellet obtained in "column buffer" (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA), another centrifugation at 9 000 g made it possible to recover a supernatant containing a high concentration of MBP-AT2. This supernatant was 15 4°C, for 3 hours at into contact, glutathione agarose beads coupled to GST protein alone after addition of NaCl so as to have 300 mM final NaCl. This preincubation step makes it possible to remove the nonspecific interactions which may exist between ATIP 20 and GA-GST. The supernatant recovered was brought into contact with the GA-GST-ATIPmouse-short or GA-GSTalone beads overnight at 4°C. After contact, the beads were rinsed three times in 20 mM Tris-HCl buffer, 300 mM NaCl, 1mM EDTA and once in "column buffer". After 25 SDS-PAGE and rinsed in beads analysing the immunoblotting with an anti-MBP antibody (New England Biolabs), a specific retention of the MBP-AT2 protein is observed on GA-GST-ATIPmouse-short beads which is not observed on the GA-GSTalone beads (Figure 10). 30

This same protocol was carried out with a plasmid expressing MBP-AT1 (C-terminal end of human AT1 receptor (Leu 297-Glu 359)); it indicates that the MBP-AT1 protein is not retained in a specific manner on the GA-GST-ATIPmouse-short beads (Figure 10).

35

These results confirm those obtained in the two-hybrid system indicating a specific and selective interaction between the protein according to the

invention and the C-terminal end of the AT2 receptor (and not AT1).

EXAMPLE 7: Modification of the transduction of the signal for the AT2 receptor in clones overexpressing the ATIPmouse-long protein

To verify that the ATIP protein interacts in vivo with the AT2 receptor, it was evaluated whether an overexpression of this protein modifies a signal induced by the AT2 receptor.

For that, a stable clone of CHO-hAT2 cells expressing the ATIPmouse-long protein (CHO-hAT2-ATIP), obtained according to the methodology described in Example 4, was used; the functional test for the activity of the AT2 receptor developed on the CHO-hAT2 clone which consists in inhibiting the phosphorylation of the IR β subunit of the insulin receptor induced by its ligand, was reproduced.

Demonstration of an inhibition by the AT2 receptor of the phosphorylation of IR β induced by insulin in CHO-hAT2 cells:

20

25

30

35

The CHO-hAT2 cells are inoculated at a density of 3×106 cells per dish having a diameter of 15 cm². They are made quiescent by 16 hours of deprivation before being treated. The treatment consists bringing into contact for 5 minutes with 15 ml of F12 medium containing insulin supplemented or otherwise with CGP42112 (selective agonist of the AT2 receptor). After treatment, the cells are solubilized in lysis buffer containing: 50 mM Hepes, pH 7.6, 1% Triton X-100, 20 mM EDTA, 30 mM sodium pyrophosphate, 30 mM sodium fluoride, 2 mM benzamidine, 1 mΜ sodium orthovanadate, 1 mM phenylmethylsulphonyl fluoride and aprotinin, pepstatin, μg/ml of antipain leupeptin. The lysates are then subjected purification on a wheatgerm lectin column, according to the protocol described in Issad, T. et al., (Eur. J. Biochem. 1995, 234, 108-115). After bringing into contact and washings, the lectin beads coupled to

Sepharose (Pharmacia) are recovered in sample buffer containing SDS and the eluted proteins are analysed in SDS-PAGE followed by immunoblotting with antiphosphotyrosine antibodies (Upstate Biotechnology, Inc.) or anti-IR β antibodies (described in Issad, T. et al., cited above).

5

10

15

20

25

30

35

The β subunit of the insulin receptor appears polypeptide of 97 kDa whose phosphorylation (visualized by revealing with an anti-phosphotyrosine antibody) increases in a dose-dependent manner with the concentration of insulin. Angiotensin II (100 nM) as well as CGP42112 (100 nM) inhibit this phosphorylation at all the insulin doses tested between 0.1 and 0.001 $\mu g/ml$ (Figure 11). By way of example, CGP42112 inhibits the phosphorylation of IR β induced by 0.01 $\mu g/ml$ by a factor of 64 \pm 4% (n=7). This result demonstrates that receptor interferes negatively with the AT2 signalling pathways for the insulin receptor at the is activation, which stage of its initial autophosphorylation. These results also provide the interconnection between first evidence of an signalling pathways for the tyrosine kinase receptors and the receptor with seven transmembrane domains which is AT2.

Reproduction of this methodology on CHO-hAT2-ATIP cells:

When this protocol is carried out on CHO-hAT2-ATIP cells, the inhibition by CGP42112 (100 nM) of the phosphorylation of the insulin receptor obtained for various doses of insulin (0.05, 0.01, 0.005, 0.001 $\mu g/ml)$ is not observed (Figure 11). This result was reproduced 3 times for each of the insulin doses taking, as positive control in each experiment, the inhibition obtained for the clone CHO-hAT2.

This therefore demonstrates that the overexpression of the ATIP protein in the CHO-hAT2 cells interferes with the AT2 receptor signalling,

which confirms the interaction in vivo of the ATIP protein with the AT2 receptor.

Another glycosylated protein, retained on a lectin column, having an apparent weight of 120 kDa, identified as being the newly cloned protein SIRP (Kharitonenkov, A. et al., Nature, 1997, 386, 181-186) is phosphorylated on tyrosine in response to insulin. The phosphorylation of this protein, as well as that of $\mbox{IR}\beta$ is inhibited in the presence of CGP42112 in the case of the clone CHO-hAT2 and is not in the case of 10 the clone CHO-hAT2-ATIP. This confirms that the ATIP protein interferes with the signalling pathways for the AT2 receptor. These results indeed show the possible value of the use of the ATIP protein for modifying signalling mediated by the AT2 receptor and 15 conditions pathological possibly compensating for associated with abnormalities in the regulation of this receptor.

As is evident from the above, the invention is not at all limited to those of its embodiments, implementations and applications which have just been described more explicitly; it encompasses, on the contrary, all the variants thereof which may occur to the specialist in the field, without departing from the framework or the scope of the present invention.

10

30

- 26 -

CLAIMS

- 1. Isolated nucleic acid fragment, encoding a protein capable of binding to the AT2 receptor, which fragment is selected from the group consisting of the sequences SEQ ID NO:1, 3, 5, 7 and 9.
 - 2. Fragment of one of the sequences according to Claim 1, comprising between 20 and 400 bp, useful as probes or as primers, for the detection of the sequences SEQ ID NO:1, 3, 5, 7 or 9, or of homologous sequences.
 - 3. Fragment according to Claim 2, characterized in that it comprises from 20 bp to 400 bp included in the sequences SEQ ID NO:1, 3, 5, 7 or 9.
- 15 4. Fragment according to Claim 2 or Claim 3, characterized in that it is selected from the group consisting of the sequences SEQ ID NO:5, SEQ ID NO:10, SEQ ID NO:11 and SEQ ID NO:12.
- 5. Transcripts, characterized in that they are complementary to the sequences according to Claim 1.
 - 6. Purified and isolated protein, which is capable of interacting with the AT2 receptor and which is selected from the group consisting of the sequences SEQ ID NO:2, 4, 6 or 8, which protein is called ATIP.
- 7. Translational product, characterized in that it is encoded by a nucleotide sequence according to Claim1.
 - 8. Antibodies, characterized in that they are directed against a protein or a protein fragment according to Claim 6 or Claim 7.
 - 9. Recombinant cloning and/or expression vector, characterized in that it comprises a nucleotide sequence according to Claim 1.
- 10. Transformed host cell, characterized in that it comprises a vector according to Claim 9.
 - 11. Transformed host cells, characterized in that they consist of a suitable yeast strain cotransformed with at least two vectors which respectively encode (i)

a so-called bait protein selected from the group consisting of a fragment containing at least SEQ ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, and a fragment containing at least the C-terminal end of the AT2 receptor, which bait protein is fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the same transcription factor and (ii) a so-called prey protein, selected from the group consisting of a fragment containing at least SEQ 10 ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, a fragment containing at least the C-terminal end of the AT2 receptor and any other polypeptide corresponding to a sequence contained in a cDNA library, which prey protein is fused with a protein 15 selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the same transcription factor, which vectors comprise, in addition, selectable markers.

- Transformed host cell according to Claim 11, 20 characterized in that it consists of a suitable yeast which cotransformed with three vectors strain respectively encode (i) a bait corresponding to a fragment containing the C-terminal end of the AT2 receptor fused with a protein selected from the group 25 consisting of the DNA-binding domain of a transcription activation domain the factor and transcription factor, (ii) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, fused with a protein selected from 30 the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the (iii) a polypeptide said transcription factor and corresponding to a sequence contained in a cDNA vectors comprise, in addition, which 35 library, selectable markers.
 - 13. Transformed host cell according to Claim 11, characterized in that it consists of a suitable yeast

cotransformed with strain two vectors which respectively encode (i) a fragment containing at least the sequence SEQ ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a polypeptide corresponding to a sequence contained in a library, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription activation domain of the factor and the transcription factor, which vectors comprise, addition, selectable markers.

5

10

35

- Transformed host cell according to Claim 11, characterized in that it consists of a suitable yeast 15 strain cotransformed with two vectors, namely (i) a vector encoding a fragment containing at least the SEQ ID NO:5 of the ATIP protein sequence according to Claim 6, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a 20 transcription factor and the activation domain of the said transcription factor and (ii) a vector encoding a fragment containing the C-terminal end of the AT2 receptor, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of 25 a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, one of the two vectors necessarily encoding a mutated protein.
- 30 15. Method for selecting proteins inhibiting ATIP protein according to Claim 6 or Claim 7-AT2 receptor interaction, which method comprises:
 - (a) cotransforming a suitable yeast strain with three vectors which respectively encode (i) a bait corresponding to a fragment containing the C-terminal end of the AT2 receptor fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the

said transcription factor, (ii) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (iii) a polypeptide corresponding to a sequence contained in a cDNA library, which vectors comprise, in addition, selectable markers,

(b) selecting the clones of cDNA library expressing a polypeptide inhibiting the AT2 receptor-ATIP protein according to Claim 6 or Claim 7 interaction, on an appropriate selective medium, and

10

20

25

30

- (c) identifying the said polypeptide.
- 15 16. Method for screening polypeptides interacting with the ATIP protein according to Claim 6 or Claim 7, which method comprises:
 - (a) cotransforming a suitable yeast strain with two vectors as defined above, namely which respectively encode (i) a fragment containing at least SEQ ID NO:5 of the ATIP protein according to Claim 6 or Claim 7, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a polypeptide corresponding to a sequence contained in a cDNA library, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, and
 - (b) selecting the clones expressing a polypeptide interacting with the ATIP protein according to Claim 6 or Claim 7, on a suitable selective medium.
- 17. Method for characterizing the domains involved 35 in the ATIP protein-AT2 receptor interaction, characterized in that it comprises:
 - (a) cotransforming a suitable yeast strain with two vectors, namely (i) a vector encoding a fragment

containing at least SEQ ID NO:5 of the ATIP protein according to Claim 6, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor and (ii) a vector encoding a fragment containing the C-terminal end of the AT2 receptor, mutated or not, fused with a protein selected from the group consisting of the DNA-binding domain of a transcription factor and the activation domain of the said transcription factor, which vectors comprise, in addition, selectable markers, one of the two vectors necessarily encoding a mutated protein, and

(b) visualizing, by selection on a suitable selective medium, the possible loss of the ATIP protein according to Claim 6 or Claim 7-AT2 receptor interaction.

10

20

25

30

35

- 18. Method for selecting substances capable of influencing the ATIP protein according to Claim 6 or Claim 7-AT2 receptor interaction, which method comprises:
- (a) bringing the ATIP protein according to Claim 6 or Claim 7, attached to a support, into contact with a fusion protein AT2 receptor-protein tag, optionally in the presence of a substance to be tested,
- (b) at least one washing of the said support thus treated with a suitable buffer, and
- (c) visualizing the possible ATIP protein according to Claim 6 or Claim 7-AT2 receptor interaction, in particular in SDS-PAGE, followed by immunoblotting with antibodies directed against the protein tag, fused with the AT2 receptor.
- 19. Method for selecting substances capable of interacting with the ATIP protein according to Claim 6 or Claim 7, characterized in that it comprises:
- (a) bringing the ATIP protein according to Claim 6 or Claim 7, attached to a support, into contact with a cell lysate,

- (b) at least one washing of the said support thus treated with a suitable buffer,
- (c) visualizing the possible protein combined with the ATIP protein, in particular in SDS-PAGE, followed by immunoblotting with appropriate antibodies, and
- (d) identifying the protein in the cell lysate interacting with the ATIP protein.
- 20. Use of the cotransformed cells according to any one of Claims 10 to 13, for the selection and screening of substances or of proteins capable of influencing the ATIP protein-AT2 receptor interaction or capable of interacting with the ATIP protein.





ORGANISATION MONDIALE DE LA PROPRIETE INT. **PCT** Bureau international

TUELLE



DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets 7:

C12N 15/12, C07K 14/47, G01N 33/68, C12Q 1/68, C12N 15/62, C07K 16/18

(11) Numéro de publication internationale:

WO 00/08148

(43) Date de publication internationale: 17 février 2000 (17.02.00)

(21) Numéro de la demande internationale:

PCT/FR99/01908

A1

(22) Date de dépôt international:

2 août 1999 (02.08.99)

(81) Etats désignés: CA, US, brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT,

(30) Données relatives à la priorité:

98/09997

4 août 1998 (04.08.98)

FR

AS-(71) Déposant (pour tous les Etats désignés sauf US): SOCIATION POUR LE DEVELOPPEMENT DE L'IMMUNOLOGIE MOLECULAIRE-ADIM [FR/FR]; 22, rue Méchain, F-75014 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (US seulement): ELBAZ, Nathalie [FR/FR]; 7, passage des Italiens, F-93170 Bagnolet (FR). NAHMIAS, Clara [FR/FR]; 4, rue Bailly, F-75003 Paris (FR). STROSBERG, Arthur, Donny [FR/FR]; 66, rue de Javel, F-75015 Paris (FR).

Publiée

Avec rapport de recherche internationale.

Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont

(74) Mandataire: CABINET ORES; 6, avenue de Messine, F-75008 Paris (FR).

(54) Title: NUCLEIC SEQUENCES CODING FOR AN AT2 INTERACTING PROTEIN INTERACTING WITH THE AT2 RECEPTOR AND THEIR APPLICATIONS

(54) Titre: SEQUENCES NUCLEIQUES CODANT POUR UNE PROTEINE (ATIP) INTERAGISSANT AVEC LE RECEPTEUR AT2 ET LEURS APPLICATIONS

(57) Abstract

The invention concerns nucleic sequences coding for a protein capable of interacting with the AT2 receptor, oligonucleotides included in said sequences, their applications as probes and for expressing said proteins, vectors useful for said expression, host cells containing said vectors, and study model of AT2 receptor. The invention also concerns said proteins and their uses. Said isolated nucleic acid fragment coding for a protein capable of binding with the AT2 receptor is selected among the group consisting of the sequences SEQ ID NO: 1, 3, 5, 7 and 9.

(57) Abrégé

Séquences nucléiques codant pour une protéine apte à interagir avec le récepteur AT2, oligonucléotides compris dans lesdites séquences, leurs applications en tant que sondes et pour l'expression desdites protéines, vecteurs utiles pour ladite expression, hôtes cellulaires contenant lesdits vecteurs, ainsi qu'un modèle d'étude du récepteur AT2. Protéines ainsi que leurs applications. Ledit fragment d'acides nucléiques isolé, codant pour une protéine apte à se lier au récepteur AT2, est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1, 3, 5, 7 et 9.

C-TERMINAL END AT2 RECEPTOR

Extrémité C-terminale récepteur AT2 160 BP DS-DNA LOCUS

Souris MOUSE ORGANISM 33 C 36 G

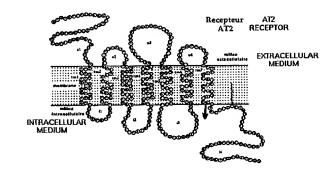
50 T

s 1 TGTGTTAATC CCTTCCTGTA TTGTTTTGTT GGAAACCGCT TCCAACAGAA CGTCCGCAGT GTGTTTAGAG TTCCCATTAC TTGGCTCCAA GGCAAGAGAG AGACTATGTC TTGCAGAAAA 121 GGCAGTTCTC TTAGAGAAAT GGACACCTTT GTGTCTTAAA ac.nucléiques NUCLEIC ACIDS

TRANSLATION INTO AMINOACIDS

Traduction en acides aminés

CVNPFLYCFV GNRFQQNVRS VFRVPITWLQ GKRETMSCRK GSSLREMDTFVS



Application No: 09/762,194 Atty Dkt No: 33339/208804

1/14

LOCUS

AT2 receptor C-terminal end

160 BP DS-DNA

ORGANISM Mouse

BASES

41 A 33 C

36 G

50 T

Nucleic acids 1 TGTGTTAATC CCTTCCTGTA TTGTTTTGTT GGAAACCGCT
TCCAACAGAA CGTCCGCAGT GTGTTTAGAG TTCCCATTAC
TTGGCTCCAA GGCAAGAGAG AGACTATGTC TTGCAGAAAA
121 GGCAGTTCTC TTAGAGAAAT GGACACCTTT GTGTCTTAAA

Translation into amino acids

CVNPFLYCFV GNRFQQNVRS VFRVPITWLQ GKRETMSCRK GS5LREMDTFVS-

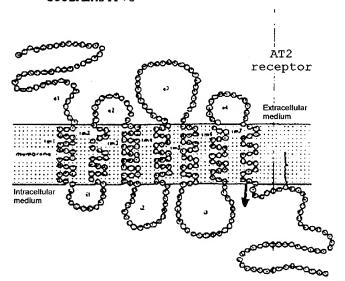


Figure 1

2/14

Codon 147 AGT AAC AAA GGT CAA AGA CAG TTG ACT GTA TCG

GAL4 DNA-binding domain

Smal

CCG GAA TTC CCG GGG ATC CGT CGA CCT... Sall

Multiple cloning site

EcoRI

BamHI

Title: NUCLEIC SEQUENCES CODING FOR AN AT2 7 7 6 21 9 4 1 Inventor(s); Elbaz, et al.
Application No: 09/762,194
Atty Dkt No: 33339/208804

3/14

	GCT	ACCC	cccc	CCCA	دحح	cccc	CCA	TCTC	GGTG	GCC.	rggc.	ertsc	CATO	: LAT	ctto	TTTT	TOTO	:TG50	7:
									GCCA										
	rcc	ccta	CGAA	GTTC	TCCC	ACTO	cer	COA	GAZ	M ATG	i crç	_ TTC	S TCT	Spc Spc	;; ДДД	F TTC	5 TCC	L TTA	9 204
	s	т	ī	н	7	R	L	7	, , sec	К	G	:	L	3	N	L	R	L CTI	27 25 8
	F	s	G	L	a	ĸ	N	ī	V GTC	I	7	Ξ	r	v	Ξ	к	s	ā	45 312
	Q	. <	N	ę	સ	s	Ŀ	c	I ATC	Ç	7	ą	Ţ	ا a	₽	Ð	v	L	63 366
	s TCC	S TCC	E GAG	R AGA	ፒ አ ሮር	CII	GXS	L TTG	A GCC	و دیم	Υ ΤλC	X AAG	T ACA	k kaa	c TGT	ςλλ Ξ	5 AGC	CYY CYY	31 4 20
	s AGT	Ç GGA	F TTC	I	L CTG	H CAC	E CTC	R AGG	Q CAG	L CTT	C.1.1	3 TCC	r Cot	G GGT	N AAC	YYC N	X AAG	; ;;;	99 474
	E GAA	A GCG	CTG	T ACA	OT:	V GTG	I ATC	Q CAG	eyc K	E CTC	L CTG	5 1 27	£ GAG	ССС	E Sag	E GAA	GCA	L CTG	117 528
	K	Q CAA	H CAC	K AAA	T ACC	ڻ CTC	S TCT	Q	E GAA	L CTT	v GTC	3 λGC	t ctc	٦ ICGG	G GGA	E GAG	L CTA	V STT	1)5 582
1	A GCT	A GCT	S TCA	S AGC	A GCC	C TGT	E GAG	K AAG	L CTA	5 5 2	Υ AAG	GCT A		ر IGCT	D GAC	i TTA	Q CAG	T ACA	133 635
	A 500	Y TAT	CYY	E E E	ŗ TTT	CI:	Q cxc	K AAA	CTA	У И	Q CAG	Ç SAG	H CAT	و حين	T ACA	GYC D	R CGG	T ACS	17 <u>1</u> 690
	E GAA	L CTG	E GAG	YYC N	R COG	í CTS	Х AAG	D GAC	C TTA	Y TAC	T ACC	A BCA	E GAG	. 555	GYQ E	77C K	CII	CY2 Ō	199 744
	5 AGC	I ATT	Y TAC	I ATT	e Gag	E GAG	A GCA	E GAA	444 K	Y TAT	K AAA	T AST	CYY O	t CTS	CAA Q	E GAG	Q CAG	ili i	207 798
<u>)</u> .	D GAC	N AAC	L TTA	N AAC	A GCC	A GCC	CYJ.	Ε GλG	T ACC	T ACT	χ λλς	<u>:</u> ::	GYC E	ATT	E G À À	GCT	S AGC	H CAC	225 352
	s TCG	S GAG	K AAG	v GTG	S GAA	L TTG	t CTG	Κ AλG	X AAG	T ACC	Y TAT	GYY CYY	ACC	S TC:	CTT	î ÇA	SÄA	A FC	243 906
	K K	K Aag	S AGC	H CAT	E GAG	M ATG	E GAG	K AAG	K AAG	5 TCA	L CTG	E Sas	D GAT	L CTG	i CTT	N AAT	E GAG	aag	261 960
2	Q CAG	E GAA	5 7CG	L CTG	GYQ	K AAA	CAA	ATC	N AAT	D GAT	L CTG	X AAG	S AGT	E GAA	AAC N	D GAT	A GCT	E TTA	279 1014
3	N N	5 2 2	R AGG	L TTC	K AAA	S TCA	£ GAG	E GAG	Q CAA	ayc K	Q CAX	: c::s	S TCA	E AGA	E GAG	Х ХХБ	GCC Y	и 245	297 1053
	S TCC	K	N AAC	P CCT	Q CAG	V GTC	M ATG	Y TAT	C.T.C	GAG	Q CAA	E Glada	CTA		S AGC	t cTG	aag	ςĉτ ——	315 1132
								ri.	חוד	_	1 1								

Figure 3.:

4/14

	V GTG																		333 1176
,		GAA	AAG			-	N AAT	_		A GÇA				, K AAG	L CTS	K AAG	R CGA	F	351 1230
4	Q CAG	Q CAG	E GAA	N AAC	E GAG	S GAG	L TTA			R CGC				ŮH ⊄AC	M ATG	A GCA	I	S TCA	359 1294
				S TCC									2 2 2 2 2		_	• • •	E GAG	-	387 1338
		v GTC		K AAG	R AGA	i CTG	5 TCC			N AAC			L CTT	L	W TGG	K AAA	CTG	eyc Cyc	405 1392
		G GGA					, CCC											F TTC	423 1445
	Q CAG	S TCC	CCC E	٦ AGG	N AAT	S TÇT	G SGT	5 TCC	F TTC	S TCC	5 AGC	CCC	S AGC	II ATC	S TCA	P CCC	R AGA	• TGA	440 1500
							CTGA												1571
							TGGA							i					1642
							.دعمم							i					1784
	CTAA	.GCAT	'AGGC	TTTC	ÇλĢ														1863

Figure 32

5/14

	cag	cgtq	acg	ggct	caga	ggca	igc : 0	CEAS	pacet	gcaç	19099	; ç 494	ttgt	A tt	agag	1944	gage	accac	= 72
	EEg	gcaa	cato	cgaa	age;	444	cgga	ages	3744	ACAC	ttgq	ccas	ecct	8323	gac	::::	tet	ctat	9 144
	cc=	ctgt	ggt	gaac	gec	::::	retet	geag	gcat	ctt	ceto	::gac	igta	eeca	ttgg	czt	gaaq	jagta	215
	LÇA	gctt	aaaa	agas	agta	cgcg	acag	teca	cgga	.382	ges		cese	iaa:	sccq	cas	cego	cccs	2 223
	aga																	G AC	
		м	_				-												17
	GCC A	AAA K	C	TTG L	1	R	N N	c 	Z Z	CTT	CC1	TCA 5	G	7	, YQQ	R R	S	ACT T	397 35
	GTT V	CTT V	TTC	CAC E	ACA T	Ç.T	<u>:</u> CYY	AAG K	AGC S	AGG R	CAA Q	K K	N N	ect ?	CGA R	AGC S	TTA L	TGT C	451 53
	ATC I	CAG Q	CCA P	CAG	ACA T	GCT A	CCC P	GAT D	GCG	CTG L) CCC	CCT	GλG £	AAA K	ACA T	CTT L	GAA E	TTG L	505 71
	ACG	CAA Q	TAT Y	ALA K	ACA	aaa K	tar c	σλλ E	AAC N	CAA Q	AGT S	GSA	III E	ATC T	c.c	CAG Q	cro	AAG K	559 39
	CAG	CTT		GCC							GY2				GIT	GTG		EXG	613
	Q	L	L	λ	¢	G	N	T	K	F	Ξ.	^		, T			I	<u> </u>	137
	H H	CTG L	CTG L	TCT S	GAG E	CSG R	GAG Z	GAA E	GCA A	CTG L	K	ς γ	K	K	λCC	CTA L	TCT 3	CAA	667 125
1	GAA	CTT	GTT V	λλC	CTC S	CGG R	GGA G	GAG E	CTA L	GTC V	ACT T	GCT A	TCA S	ACC T	ACC T	TGT	GAG E	AAA K	721 163
	TTA	GYY 3	مند ×	GCC A	AGG R	AAT N	GAG	TTA L	CAA	ACA T	GŢG	TAT Y	GAA E	GCA À	TTC T	GTC	CAG	CAG	775 151
	CAC	CAG	GCT	GAA	ممد	ACA	GAA	CGA	GAG	AAT	CGG	CTT	AAA	GAG		TAC	ACC	AGG	329
	ä	Q	λ	E	ĸ	T	Ξ	R	2	N	2.	Ŀ	K	Ξ	7	Y	Ţ	R	179
	GAG E	TAT Y	GAA E	aag K	CT:	CGG R	GAC D	ACT T	TAC Y	ATT	Ē CYV	E CVY	γ GEX	GAG E	AAG K	TAC Y	AAA K	ATS H	883 197
	CAA		CAA		CAG	TTT		AAC		AAT		CAT	GAA			AAG	TTG		33°
Ļ	Q	Ľ.	•	E	Q		<u> </u>	N	Ļ	N	٨	#	Ε	Т	5	K		<u> </u>	215
	ATT I	Gλλ	GCT A	AGC S	CAC H	TCA S	ςλG 3	aaa K	CTT S	GAA E	TTG 1	CTA L	AAG K	AAG X	GCC A	TAT	GAA E	A GCI	991 233
	TCC	CII	TCA S	GAA E	ATT I	K WG	***	GGC G	TAS Y	E GYY	ATA I	E E E	AAG K	XXX K	TCG S	CTT L	E Gyy	GAT D	1345 251
	TTA	CT:	TCT	GAG		CAG				GAG			ATC		GAT			AGT	1095
	L	į.	\$	E	К	0		\$	Ľ	E	<u> </u>	2		N		L	K	5	269
3	GAA E	AAT N	GAT D	GCT Å	TTA L	aat N	E E	AAA K	tTG L	×	TCA 5	SAA E	CAA Z	CAA Q	ж Х	AGA R	AGA R	GCA A	1153 267
	AGA R	GAA E	AAA K	GCY V	ж .	TTG	AAA K	AAT N	CCT	CAG Q	ATC £	ATC M	TAT Y	CTA L	GAA E	CAG 2	GλG E	TTA L	1207 305
																		_	

Figure 41

Inventor(s): Elbaz, et al. Application No: 09/762,194 Atty Dkt No: 33339/208804

GAA AGC CTG AAA GCT GTG TTA GAG ATC AAG AAT GAG AAA CTG CAT CAA CAG GAC E S L K A V L E I K N E K L H O C S CTG TGG AAA CTG CAC AAT GGG GAC CTG TGT AGC CCC AAA AGA TCC CCC ACA TCC L W K L H N G D L C S P K % S P T S 1535 TOO GOO ATC COT TTG CAG TOA GOA AGG AAT TOO GGO TTO TTO COT AGG COO AGG ATT TCA CCC AGA TGA cacgtocccaaagtetacagactetttgaaagdattttgaageaggtotgc 1551 aggactgaccccaaggaggaecgtgggcacaagaggtatatcaccacgtgtgtatcaccgtaggtaaccgg 1723 agogttaccacoggeggaatogagettotgagactggaagttttggaggaagac\$ttttgcctccgtttaaaag 1795 actoctccamamangatttamamangatttcqqcatcqacacqgacqttqttqcacamagcacttamaga 1967 acgagageacettgtteacifeetetteacetaageataaggggaaaaactetcagggetetattaagatt 1939 cataacottogtaatgttcttcaccacagacaccttcttgtgagtttttagttfgactgtgggggtgggggg 2011 azotgaatatazotggatasgsactaatgaataataatcaatcaatsagsatåtacattttagscoaaagco 2155 adagaagaaaaagcaatagttgcttgaattatgatcatctactactactatttttttcagccctqtaacagggt 2227 agggagagggtataacaggaagagctt:gacttgtccccgtctatacattcctctgtgtatcttttgggggtaac 2299 etettggcagttttttagtgtttagccatgtcagllgaaactagattttttgtagattttttttactacttaccca 2371 Egggagoctaacactaccctgtaactcatttccccaggctacgtgtalatgtagaaccctaactcctcctata 2443 contribution of the second sec acquettteccagagecccagagecagettatettettaggtgetgaesttaetteteaaataaactaaageet 2731 ggatttgatattataaattttgggaaattttagaatskagatttaagghagttattaggttattaa 2803 ccaagaaaggcaggacccagcgcccaccgacgcagtccccgacctaagaabacagcccggcaggacagc 2975 gotcatotetecagttaccetetaaggagteccotttgtotttgggaAagtagcagaatggtccgcttcttcc 2947 catgagtggaasatgtggettgtccaastetectecaeggttgcallipagtticcttccaaaacttattacr 1019

Figure 4.2

Title: NUCLEIC SEQUENCES CODING FOR AN AT2. CISTED 9476121194

Inventor(s): Elbaz, et al. Application No: 09/762,194 Atty Dkt No: 33339/208804

7/14

Figure 4.3

Inventor(s): Elbaz, et al. Application No: 09/762,194 Atty Dkt No: 33339/208804

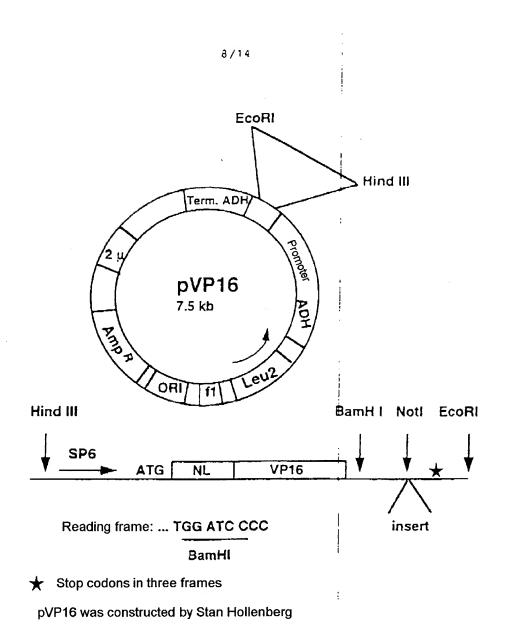


Figure 5

Title: NUCLEIC SEQUENCES CODING FOR AN AT2... Inventor(s): Elbaz, et al.
Application No: 09/762,194
Atty Dkt No: 33339/208804

9/14

ATG	GAT	J -	CGA	AAG	•
CAT GGT ATG	CGG	GGA TCC	AG .	ACA AAG CCC GAA AGG	TAG
CAT	GGT	766	GGA	GAA	AAC
CAT	ATG	CGA	CAT GGA	ວລວ	CTG AGT TGG CTG CCG CTG AGC AAT AAC
CAT	GCT AGC ATG ACT GGT GGA CAG CAA ATG	CTG TAC GAC GAT GAC GAT AAG GAT CGA	GAG CTC GAG ATC TGC AGC TGG TAC	AAG	AGC
CAT	CAG	AAG	TGG	ACA	CTG
CAT	GGA	GAT	AGC	CTG CTA	೮೦೦
CAT	GGT	GAC	TGC	CTG	CCA
TCT	ACT	GAT	ATC	TTG ATC CGG	CTG
GGT	ATG	GAC	GAG	ATC	TGG
000	AGC	TAC	CTC	TTG	AGT
98. ATG CGG GGT TCT CAT CAT CAT CAT	GCT	CTG	GAG	AGC	CTG
98.	134	170	206	242	278

6 histidines

Title: NUCLEIC SEQUENCES CODING FOR AN AT2... Inventor(s): Elbaz, et al. Application No: 09/762,194 Atty Dkt No: 33339/208804

10/14

ATG GGT CCG GAA CAG AAA CTG ATC TCT GAA GAA CAC CTG GGA TCC GGA ATT CTA GA דאכ ככא, מככ כדד מדכ דדד מאכ דאם אמא כדני כידי כדם סאכן ככד אסם ככד דאא מאד כז Met gly proglu gin lys leu ile ser glu glu asp leu gly ser gly ile leu Tag Myc

Inventor(s): Elbaz, et al.

Application No: 09/762,194

Atty Dkt No: 33339/208804

14/14

CHO-hAT2

Lectin column

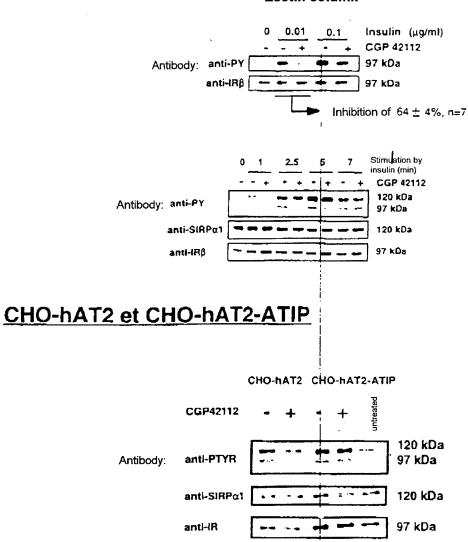
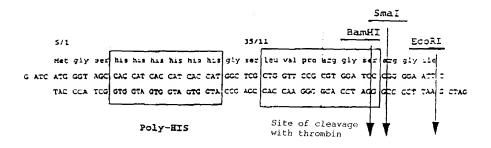


Figure 11

Title: NUCLEIC SEQUENCES CODING FOR AN AT2...

Inventor(s): Elbaz, et al. Application No: 09/762,194 Atty Dkt No: 33339/208804

11/14



pBacPAKI-poly HIS -> Graphic Map

PolyHIS insertion into pBackpack in BamHI(CACCAT) 1270-1287

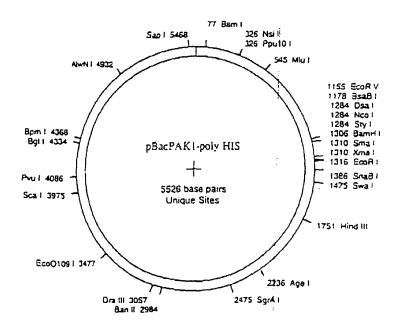


Figure 8

Title: NUCLEIC SEQUENCES CODING FOR AN AT2 USSE 1047.6.219 41.
Inventor(s): Elbaz, et al.
Application No: 09/762,194
Atty Dkt No: 33339/208804

12/14

Tissues:

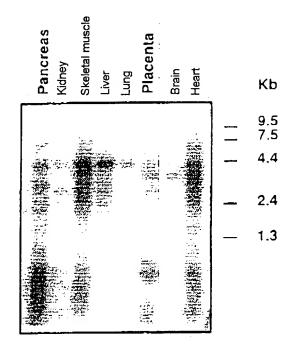


Figure 9

Title: NUCLEIC SEQUENCES CODING FOR AN AT2 The Transfer of the Control of the Con

13/14 GST-ATIP ▲ MBP-AT2 **GST**alone MBP-AT2 MBPv MBP-AT1 Supernatants: 111 I 1 **GST-ATIP GST**alone KDa 48 40 33 anti-MBP Beads: Antibodies

Figure 10

Declaration and Power of Attorney for Patent Application Déclaration et Pouvoirs pour Demande de Brevet French Language Declaration

En tant l'inventeur nommé ci-après, je déclare par le présent acte que :

Mon domicile, mon adresse postale et ma nationalité sont ceux figurant ci-dessous à côté de mon nom.

Je crois être le premier inventeur original et unique (si un scul nom est mentionné cidessous), ou l'un des premiers co-inventeurs originaux (si plusieurs noms sont mentionnés ci-dessous) de l'objet revendiqué, pour lequel une demande de brevet a été déposée concernant l'invention intitulée

et dont la description est fournie ci-joint à moins

ţ

ci-joint

a été déposée le

sous le numéro de demande des Etats-Unis ou le numéro de demande international PCT

et modifiée le

(le cas échéant).

Je déclare par le présent acte avoir passé en revue et compris le contenu de la description ci-dessus, revendications comprises, telles que modifiées par toute modification dont il aura été fait références ci-dessus.

Je reconnais devoir divulguer toute information pertinente à la brevetabilité, comme défini dans le Titre 37, § 1.56 du Code fédéral des réglementations.

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed an for which a patent is sought on the invention entitled

Nucleic sequences coding for an AT2 interacting protein interacting with the AT2 receptor and their applications

the specification of which:

is attached hereto.

was filed on

as United States Application Number or PCT International Application Number. PCT/FR99/01908 filed on August 2, 1999

and was amended on

(if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations,§ 1.56.

French Language Declaration

Je revendique par le présent acte avoir la priorité étrangère, en vertu du Titre 35, § 119(a)-(d) ou § 365(b) du Code des Etats-Unis, sur toute demande étrangère de brevet ou certificat d'inventeur ou, en vertu du Titre 35, § 365(s) du même Code, sur toute demande internationale PCT désignant au moins un pays autre que les Etats-Unis et figurant ci-dessous et, en cochant la case, j'ai aussi indiqué ci-dessous toute demande étrangère de brevet, tout certificat d'inventeur ou toute demande internationale PCT ayant date de dépôt précédant celle de la domande à propos de laquelle une priorité est revendiquée.

Prior Forcign application(s)
Demande(s) de brevet antérieure(s) dans un autre pays.
FR 98 08600 FRANCE

(Number) (Country)
(Numéro) (Paye)
(Number) (Country)

(Numéro) (Pays)

Je revendique par le présent acte tout bénéfice, en vertu du Titre 35, § 119(e) du Code des Etats-Unis, de toute demande de brevet provisoire effectuée aux Etats-Unis et figurant ci-dessous.

(Application No.) (Filing Date) (N° de demande) (Date de dépôt)

Je revendique par le présent acte tout bénéfice, en vertu du Titre 35, § 120 du Code des Elats-Unis, de toute demande de brevet effectuée aux Etats-Unis, ou en vertu du Titre 35, § 365© du même Code, de toute demande internationale PCT désignant les Etats-Unis et figurant ci-dessous et, dans la mesure où l'objet de chacune des revendications de cette demande de brevet n'est pas divulgué dans la demande antérieure américaine ou internationale PCT, en vertu des dispositions du premier paragraphe du Titre 35, § 112 du code des Etats-Unis, je reconnais devoir divulguer toute information pertinente à la brevetabilité, comme défini dans le Titre 37, § 1.56 du Code fédéral des réglementations, dont j'ai pu disposer entre la date de dépôt de la demande antérieure et la date de dépôt de la demande antérieure et la date de depôt de la demande nationale ou internationale PCT de la présente demande:

(Application No.) (Filing Date) (N° de demande) (Date de dépôt)

(Application No.) (Filing Date) (N° de demande) (Date de dépôt)

Je déclare que par le présent acte que toute déclaration ci-incluse est, à ma connaissance, véridique et que toute déclaration formulée à partir de renseignements ou de suppositions est tenue pour véridique ;et de plus, que toutes ces déclarations ont été formulées en sachant que toute fausse déclaration volontaire ou son équivalent est passible d'une amende ou d'une incarcération, ou des deux, en vertu de la section 1001 du Titre 18 du Code de Etats-Unis, et que de telles déclarations volontairement fausses risquent de compromettre la validité de la demande de brevet ou du brevet délivré à partir de celle-ci.

I hereby claim foreign priority under Title 35, United States Code, § 119(a)-(d) or § 365(b) of any foreign application(a) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below, and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Priority claimed
Droit de priorité
reyendiqué

August 4, 1998

(Day/Month/Year Filed)
(Jour/Mois/Anné de dépôt)

(Day/Month/Year Filed)
(Day/Month/Year Filed)
(Day/Month/Year Filed)
(Jour/Mois/Anné de dépôt)

(Day/Month/Year Filed)
(Jour/Mois/Anné de dépôt)

Oui

Non

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below.

(Application No.) (Filing Date) (N° de demande) (Date de dépôt)

I hereby claim the benefit under. Title 35, United States Code, § 120 of any United States application(s), or § 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

(Status) (patented, pending, abandoned) (Statut) (breveté, en cours d'examen, abandonné)

(Status) (patented, pending, abandoned) (Statut) (breveté, en cours d'examen, abandonné)

I hebery declare that all statements made herein of my own knowledge are true and that all statements made on information and bolief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

French Language Declaration

POUVOIRS: En tant que l'inventeur cité, je désigne par la présente l'(les) avocats(s) et/ou agent(s) suivant(s) pour qu'ils poursuive(nt) la procédure de cette demande de brevet et traite(nt) toute affaire s'y rapportant avec l'Office des brevets et des marquees: (mentionner le nom et le numéro d'enregistrement). POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to persecute this application and transact all bussiness in the Patent and Trademark Office connected therewith: (list name and number) registration

All practitioners associated with **CUSTOMER NUMBER 000826**

RAYMOND O. LINKER, JR. Registration No. 26,419

Adresser toute	correspondance	A٠

Send Correspondence to:

ALSTON & BIRD LLP
Bank of America Plaza
101 South Tyron Street, Suite 4000
CHARLOTTE, NC 28280-4000 U.S.A.

Adresser tout appel téléphonique à : (nom et numéro de téléphone)

Direct Telephone calls to: (name and telephone number)

(704) 444-1000

O	Nom complete de l'unique ou premier inventeur ELBAZ Nathalle	Full name of sole or first inventor	
, ,	Signature l'inventeur Date	Inventor's signature	Date
	Domicile 93170 BAGNOLET (FRANCE) FP	Residence	
	Nationalité Française	Citizenship	
	Adresse Postale 7, Passage des Italiens 93170 BAGNOLET FRANCE	Post Office Address	
200	Nom complete du second co-inventeur, le cas echeant NAHMIAS Clara	Full name of accord joint inventor, if any	
XX	Signature de l'inventeur Date 6/04/01	Second inventor's signature	Date
	Domicile 75003 Paris (ERANCE)	Residence	
	Nationalité Prançaise	Citizenship	
	Adresse Postale 4, Rue Ballly 75003 PARIS FRANCE	Post Office Address	
	Let our de Turenne - 20002 Pais		

(Fournir les mêmes renseignements et la signature de tout coinventeur supplémentaire.)

(Suppply similar information and signature for third and subsequent joint inventors.)

Page 3 of 4

French Language Declaration

	Nom complete du troisième co-inventeur, le cas échéant		Full name of third joint inventor, if any	
ann	STROSBERG Arthur, Donny			Date
300	Signature de l'inventeur	Date	Third inventor's signature	17ate
	Domicile 7:5015 Paus (FRANCE)	<u> </u>	Residence	
	Nationalité Française		Citizenship	
	Adresse Postale 66, Rue de Javel 75015 PARIS FRANCE		Post Office Address	
	Nom complete du quatrième co-inventeur, le cas echeant		Full name of fourth joint inventor, if any	
	Signature de l'Inventeur	Date	Fourth inventor's signature	Date
	Domicile		Residence	
	Nationalité		Citizenship	
	Adresse Postale		Post Office Address	
	Nom complete du cinquième co-inventeur, le cas ocheant		Full name of fifth joint inventor, if any	
	Signature de l'inventeur	Date	Fifth inventor's signature	Date
	Domicile		Residence	
	Nationalité		Citizenship	
<u>.</u>	Adresse Postale		Post Office Address	
	Nom complete du sixième co-inventeur, le cas echeant		Full name of sixth joint inventor, if any	
Ì	Signature de l'inventeur	Date	Sixth inventor's aignature	Date
	Domicile		Residence	
	Nationalité		Citizonship	
	Adresse Postalo		Post Office Address	
		_		

(Fournir les mêmes renseignements et la signature de tout co-inventeur supplémentaire.)

Supply similar information and signature for third and subsequent joint inventors.)

Rec'd PCT/PTO 19 APR 2001 09/762194

SEQUENCE LISTING

<110> Elbaz, Nathalie Nahmias, Clara Strosberg, Arthur Donny

<120> NUCLEIC SEQUENCES ENCODING AN AT2
 RECEPTOR-INTERACTING PROTEIN (ATIP) AND THEIR APPLICATIONS

<130> 33339/208804	,
<140> US 09/762,194 <141>	
<150> PCT/FR99/01908 <151> 1999-08-02	· .
<150> FR 98/09997 <151> 1998-08-04	
<160> 12	
<170> FastSEQ for Windows Version 4.0	
<210> 1 <211> 1803 <212> DNA <213> Mus musculus	
<220> <221> CDS <222> (178)(1500)	
<400> 1 gctaccccc ccccacgcac cccccaatct gggtg ttttctctgg ctgtatctct tggcctggaa gaacc tgatggtccc tggaaaagct gcttcccctg cgaag	ccgag ttgccaagag acacagtatg 120
ctg ttg tct ccc aaa ttc tcc tta tcc acc Leu Leu Ser Pro Lys Phe Ser Leu Ser Th	
gcc aaa gga ctg ctt cga aac ctc cgg ct Ala Lys Gly Leu Leu Arg Asn Leu Arg Le 20 25	
aac act gtc att ttc cac aca gtt gaa aa Asn Thr Val Ile Phe His Thr Val Glu Ly	

	35					40					45					
agg Arg 50	agc Ser	ctg Leu	tgc Cys	atc Ile	cag Gln 55	acc Thr	cag Gln	aca Thr	gct Ala	cca Pro 60	gat Asp	gtg Val	ctg Leu	tcc Ser	tcc Ser 65	372
gag Glu	aga Arg	acg Thr	ctt Leu	gag Glu 70	ttg Leu	gcc Ala	caa Gln	tac Tyr	aag Lys 75	aca Thr	aaa Lys	tgt Cys	gaa Glu	agc Ser 80	caa Gln	420
						ctc Leu										468
						gtt Val										516
						cac His 120										564
ctc Leu 130	cgg Arg	gga Gly	gag Glu	cta Leu	gtt Val 135	gct Ala	gct Ala	tca Ser	agc Ser	gcc Ala 140	tgt Cys	gag Glu	aag Lys	cta Leu	gaa Glu 145	612
aag Lys	gct Ala	agg Arg	gct Ala	gac Asp 150	tta Leu	cag Gln	aca Thr	gcg Ala	tat Tyr 155	caa Gln	gaa Glu	ttt Phe	gtc Val	cag Gln 160	aaa Lys	660
						aca Thr										708
						gag Glu										756
			Glu	Lys	Tyr	aaa Lys 200	Thr	Gln								804
tta Leu 210	aac Asn	gcc Ala	gcc Ala	cat His	gag Glu 215	acc Thr	act Thr	aag Lys	ctt Leu	gag Glu 220	att Ile	gaa Glu	gct Ala	agc Ser	cac His 225	852
tcg Ser	gag Glu	aag Lys	gtg Val	gaa Glu 230	ttg Leu	ctg Leu	aag Lys	aag Lys	acc Thr 235	tat Tyr	gaa Glu	acc Thr	tcc Ser	ctt Leu 240	tca Ser	900
gaa Glu	atc Ile	aag Lys	aag Lys 245	agc Ser	cat His	gag Glu	atg Met	gag Glu 250	aag Lys	aag Lys	tca Ser	ctg Leu	gag Glu 255	gat Asp	ctg Leu	948

											atc Ile					996
											tca Ser 285					1044
											cct Pro					1092
					Glu						tta Leu					1140
											aag Lys					1188
											aag Lys					1236
											cac His 365					1284
											gag Glu					1332
											aac Asn					1380
											aag Lys					1428
											tct Ser					1476
_		agc Ser					tga *	cggo	ttct	ga a	egca	iggag	ja ct	ctct	gaag	1530
ctct aata	ctgg igaga	gaa t agc t	atco	ccaq ataga	gg at aa ga	atco	ggag taac	ago ttg	agco gtco	gcc aaa	aacc agcc	gtat	ca g	gctac naaaa	egttag egtacg acagat egttet	1590 1650 1710 1770

<210> 2 <211> 440 <212> PRT <213> Mus musculus Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Val Arg Leu Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Leu Arg Lys Asn Thr Val Ile Phe His Thr Val Glu Lys Gly Arg Gln Lys Asn 40 Pro Arg Ser Leu Cys Ile Gln Thr Gln Thr Ala Pro Asp Val Leu Ser 55 Ser Glu Arg Thr Leu Glu Leu Ala Gln Tyr Lys Thr Lys Cys Glu Ser 70 Gln Ser Gly Phe Ile Leu His Leu Arg Gln Leu Leu Ser Arg Gly Asn 90 Asn Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu 105 Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val 120 Ser Leu Arg Gly Glu Leu Val Ala Ala Ser Ser Ala Cys Glu Lys Leu 135 Glu Lys Ala Arg Ala Asp Leu Gln Thr Ala Tyr Gln Glu Phe Val Gln 155 Lys Leu Asn Gln Gln His Gln Thr Asp Arg Thr Glu Leu Glu Asn Arg 165 Leu Lys Asp Leu Tyr Thr Ala Glu Cys Glu Lys Leu Gln Ser Ile Tyr 185 Ile Glu Glu Ala Glu Lys Tyr Lys Thr Gln Leu Gln Glu Gln Phe Asp 205 200 Asn Leu Asn Ala Ala His Glu Thr Thr Lys Leu Glu Ile Glu Ala Ser 220 215 His Ser Glu Lys Val Glu Leu Leu Lys Lys Thr Tyr Glu Thr Ser Leu 230 235 Ser Glu Ile Lys Lys Ser His Glu Met Glu Lys Lys Ser Leu Glu Asp 245 250 Leu Leu Asn Glu Lys Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu 265 Lys Ser Glu Asn Asp Ala Leu Asn Glu Arg Leu Lys Ser Glu Glu Gln Lys Gln Leu Ser Arg Glu Lys Ala Asn Ser Lys Asn Pro Gln Val Met 295 Tyr Leu Glu Glu Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys 315 310 Asn Glu Lys Leu His Gln Gln Asp Met Lys Leu Met Lys Met Glu Lys 330 Leu Val Asp Asn Asn Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln 345 Gln Glu Asn Glu Glu Leu Lys Ala Arg Met Asp Lys His Met Ala Ile

		355			•		360					365				
Ser	Arg 370	Gln	Leu	Ser	Thr	Glu 375	Gln	Ala	Ala	Leu	Gln 380	Glu	Ser	Leu	Glu	
Lys 385	Glu	Ser	Lys	Val	Asn 390	Lys	Arg	Leu	Ser	Met 395	Glu	Asn	Glu	Glu	Leu 400	
Leu	Trp	Lys	Leu	His 405	Asn	Gly	Asp	Leu	Cys 410	Ser	Pro	Lys	Arg	Ser 415	Pro	
Thr	Ser	Ser	Ala 420	Ile	Pro	Phe	Gln	Ser 425	Pro	Arg	Asn	Ser	Gly 430	Ser	Phe	
Ser	Ser	Pro 435	Ser	Ile	Ser	Pro	Arg 440									
<212 <212	0 > 3 L > 13 2 > Di		ıggul	lue												
<220		15 III	iscu:	Lub								*1				
<22	L> CI	os 1)	. (132	23)												
	0> 3 cta	tta	tct	ccc	aaa	ttc	tcc	tta	tcc	acc	atc	cac	atc	cgc	cta	4.8
														Arg 15		
	_			_		_						_		ctc Leu		96
														aag Lys		144
														ctg Leu		192
		_	_			_								gaa Glu		240
														ggt Gly 95		288
	_		_		_									tct Ser		336
														ctt Leu		384

agc Ser	ctc Leu 130	cgg Arg	gga Gly	gag Glu	cta Leu	gtt Val 135	gct Ala	gct Ala	tca Ser	agc Ser	gcc Ala 140	tgt Cys	gag Glu	aag Lys	cta Leu	432	
gaa Glu 145	aag Lys	gct Ala	agg Arg	gct Ala	gac Asp 150	tta Leu	cag Gln	aca Thr	gcg Ala	tat Tyr 155	caa Gln	gaa Glu	ttt Phe	gtc Val	cag Gln 160	480	
Lys	Leu	Asn	Gln	cag Gln 165	His	Gln	Thr	Asp	Arg 170	Thr	Glu	Leu	Glu	Asn 175	Arg	528	
Leu	Lys	Asp	Leu 180	tac Tyr	Thr	Ala	Glu	Cys 185	Glu	Lys	Leu	Gln	Ser 190	Ile	Tyr	576	
Ile	Glu	Glu 195	Ala	gaa Glu	Lys	Tyr	Lys 200	Thr	Gln	Leu	Gln	Glu 205	Gln	Phe	Asp	624	
Asn	Leu 210	Asn	Ala	gcc Ala	His	Glu 215	Thr	Thr	Lys	Leu	Glu 220	Ile	Glu	Ala	Ser	672	
His 225	Ser	Glu	Lys	gtg Val	Glu 230	Leu	Leu	Lys	Lys	Thr 235	Tyr	Glu	Thr	Ser	Leu 240	720	
Ser	Glu	Ile	Lys	aag Lys 245	Ser	His	Glu	Met	Glu 250	Lys	Lys	Ser	Leu	Glu 255	Asp	768	
Leu	Leu	Asn	Glu 260	aag Lys	Gln	Glu	Ser	Leu 265	Glu	Lys	Gln	Ile	Asn 270	Asp	Leu	816	
aag Lys	agt Ser	gaa Glu 275	aac Asn	gat Asp	gct Ala	tta Leu	aac Asn 280	gaa Glu	agg Arg	ttg Leu	aaa Lys	tca Ser 285	gag Glu	gag Glu	caa Gln	864	
Lys	Gln 290	Leu	Ser	aga Arg	Glu	Lys 295	Ala	Asn	Ser	Lys	Asn 300	Pro	Gln	Val	Met	912	
Tyr 305	Leu	Glu	Gln	gaa Glu	Leu 310	Glu	Ser	Leu	Lys	Ala 315	Val	Leu	Glu	Ile	Lys 320	960	
aat Asn	gag Glu	aag Lys	ctg Leu	cac His 325	cag Gln	cag Gln	gac Asp	atg Met	aag Lys 330	cta Leu	atg Met	aag Lys	atg Met	gaa Glu 335	aag Lys	1008	

Leu Val			aac Asn												1056
cag gaa Gln Glu															1104
tca agg Ser Arg 370					_		_								1152
aag gag Lys Glu 385															1200
ctg tgg Leu Trp															1248
acc tcc Thr Ser															1296
tcc agc Ser Ser		_	_			_	tga *								1323
<210> 4 <211> 44 <212> PF <213> Mu	T	ıscu]	lus											Z.	
<211> 44 <212> PF <213> Mu <400> 4	RT ıs mu				Dha	Com	Lou	Cor	The	T]o	Hig	val	hra		
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1	RT ıs mu Leu	Ser	Pro 5					10					15		
<211> 44 <212> PF <213> Mu <400> 4 Met Leu	RT ıs mu Leu	Ser	Pro 5					10					15		
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1	RT is mu Leu Lys	Ser Gly 20	Pro 5 Leu	Leu	Arg	Asn	Leu 25	10 Arg	Leu	Pro	Ser	Gly 30	15 Leu	Arg	
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1 Thr Ala	ET IS MU Leu Lys Thr 35	Ser Gly 20 Val	Pro 5 Leu Ile	Leu Phe	Arg His	Asn Thr 40	Leu 25 Val	10 Arg Glu	Leu Lys	Pro Gly	Ser Arg 45	Gly 30 Gln	15 Leu Lys	Arg Asn	
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1 Thr Ala Lys Asn Pro Arg 50 Ser Glu	Leu Lys Thr 35 Ser	Ser Gly 20 Val Leu	Pro 5 Leu Ile Cys	Leu Phe Ile Glu	Arg His Gln 55	Asn Thr 40 Thr	Leu 25 Val Gln	10 Arg Glu Thr	Leu Lys Ala Lys	Pro Gly Pro 60	Ser Arg 45 Asp	Gly 30 Gln Val	15 Leu Lys Leu	Arg Asn Ser Ser	
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1 Thr Ala Lys Asn Pro Arg 50	Leu Lys Thr 35 Ser	Ser Gly 20 Val Leu Thr	Pro 5 Leu Ile Cys Leu	Leu Phe Ile Glu 70	Arg His Gln 55 Leu	Asn Thr 40 Thr	Leu 25 Val Gln	10 Arg Glu Thr Tyr	Leu Lys Ala Lys 75	Pro Gly Pro 60 Thr	Ser Arg 45 Asp Lys	Gly 30 Gln Val Cys	15 Leu Lys Leu Glu Gly	Arg Asn Ser Ser 80	
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1 Thr Ala Lys Asn Pro Arg 50 Ser Glu 65	Leu Lys Thr 35 Ser Arg	Ser Gly 20 Val Leu Thr Phe Glu	Pro 5 Leu Ile Cys Leu Ile 85	Leu Phe Ile Glu 70 Leu	Arg His Gln 55 Leu His	Asn Thr 40 Thr Ala Leu	Leu 25 Val Gln Gln Arg	10 Arg Glu Thr Tyr Gln 90	Leu Lys Ala Lys 75 Leu	Pro Gly Pro 60 Thr	Ser Arg 45 Asp Lys Ser	Gly 30 Gln Val Cys Arg	15 Leu Lys Leu Glu Gly 95	Arg Asn Ser Ser 80 Asn	
<211> 44 <212> PF <213> Mu <400> 4 Met Leu 1 Thr Ala Lys Asn Pro Arg 50 Ser Glu 65 Gln Ser	Leu Lys Thr 35 Ser Arg Gly	Ser Gly 20 Val Leu Thr Phe Glu 100	Pro 5 Leu Ile Cys Leu Ile 85 Ala	Leu Phe Ile Glu 70 Leu Leu	Arg His Gln 55 Leu His	Asn Thr 40 Thr Ala Leu Val	Leu 25 Val Gln Gln Arg Val 105	10 Arg Glu Thr Tyr Gln 90 Ile	Leu Lys Ala Lys 75 Leu Gln	Pro Gly Pro 60 Thr Leu	Ser Arg 45 Asp Lys Ser Leu	Gly 30 Gln Val Cys Arg Leu 110	15 Leu Lys Leu Glu Gly 95 Ser	Arg Asn Ser Ser 80 Asn	

```
Glu Lys Ala Arg Ala Asp Leu Gln Thr Ala Tyr Gln Glu Phe Val Gln
                   150
Lys Leu Asn Gln Gln His Gln Thr Asp Arg Thr Glu Leu Glu Asn Arg
                                   170
               165
Leu Lys Asp Leu Tyr Thr Ala Glu Cys Glu Lys Leu Gln Ser Ile Tyr
                              185
Ile Glu Glu Ala Glu Lys Tyr Lys Thr Gln Leu Gln Glu Gln Phe Asp
                           200
Asn Leu Asn Ala Ala His Glu Thr Thr Lys Leu Glu Ile Glu Ala Ser
                                           220
                       215
His Ser Glu Lys Val Glu Leu Leu Lys Lys Thr Tyr Glu Thr Ser Leu
                                       235
                  230
Ser Glu Ile Lys Lys Ser His Glu Met Glu Lys Lys Ser Leu Glu Asp
                                  250
               245
Leu Leu Asn Glu Lys Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu
                               265
Lys Ser Glu Asn Asp Ala Leu Asn Glu Arg Leu Lys Ser Glu Glu Gln
                           280
Lys Gln Leu Ser Arg Glu Lys Ala Asn Ser Lys Asn Pro Gln Val Met
                       295
Tyr Leu Glu Gln Glu Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys
                                       315
Asn Glu Lys Leu His Gln Gln Asp Met Lys Leu Met Lys Met Glu Lys
               325
Leu Val Asp Asn Asn Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln
                               345
Gln Glu Asn Glu Glu Leu Lys Ala Arg Met Asp Lys His Met Ala Ile
                           360
Ser Arg Gln Leu Ser Thr Glu Gln Ala Ala Leu Gln Glu Ser Leu Glu
                       375
Lys Glu Ser Lys Val Asn Lys Arg Leu Ser Met Glu Asn Glu Glu Leu
                                       395
                   390
Leu Trp Lys Leu His Asn Gly Asp Leu Cys Ser Pro Lys Arg Ser Pro
                                  410
               405
Thr Ser Ser Ala Ile Pro Phe Gln Ser Pro Arg Asn Ser Gly Ser Phe
                               425
           420
Ser Ser Pro Ser Ile Ser Pro Arg
<210> 5
<211> 354
<212> DNA
<213> Artificial Sequence
<220>
<221> CDS
<222> (1)...(354)
<223> Insert identified by two-hybrid screening of a M.
      musculus foetal cDNA library
<223> Insert identified by two-hybrid screening of a M.
```

musculus foetal cDNA library

<400)> 5																
cat His 1	cag Gln	aca Thr	gac Asp	cgg Arg 5	acg Thr	gaa Glu	ctg Leu	gag Glu	aac Asn 10	cgg Arg	ctg Leu	aag Lys	gac Asp	Leu 15	tac Tyr		48
acc Thr	gca Ala	gag Glu	tgt Cys 20	gag Glu	aag Lys	ctt Leu	cag Gln	agc Ser 25	att Ile	tac Tyr	att Ile	gag Glu	gag Glu 30	gca Ala	gaa Glu		96
aaa Lys	tat Tyr	aaa Lys 35	act Thr	caa Gln	ctg Leu	caa Gln	gag Glu 40	cag Gln	ttt Phe	gac Asp	aac Asn	tta Leu 45	aac Asn	gcc Ala	gcc Ala	1	44
cat His	gag Glu 50	acc Thr	act Thr	aag Lys	ctt Leu	gag Glu 55	att Ile	gaa Glu	gct Ala	agc Ser	cac His 60	tcg Ser	gag Glu	aag Lys	gtg Val	1	92
gaa Glu 65	ttg Leu	ctg Leu	aag Lys	aag Lys	acc Thr 70	tat Tyr	gaa Glu	acc Thr	tcc Ser	ctt Leu 75	tca Ser	gaa Glu	atc Ile	aag Lys	aag Lys 80	2	40
agc Ser	cat His	gag Glu	atg Met	gag Glu 85	aag Lys	aag Lys	tca Ser	ctg Leu	gag Glu 90	gat Asp	ctg Leu	ctt Leu	aat Asn	gag Glu 95	aag Lys	2	88
cag Gln	gaa Glu	tcg Ser	ctg Leu 100	gag Glu	aaa Lys	caa Gln	atc Ile	aat Asn 105	gat Asp	ctg Leu	aag Lys	agt Ser	gaa Glu 110	aac Asn	gat Asp	3	36
_			gaa Glu													3	54
<21 <21	0> 6 1> 1: 2> Pl 3> A:	RT	icial	l Sed	quen	ce											
<22 <22	3> I		t ide							orid	scre	eeni	ng o:	£			
	0> 6 Gln	Thr	Asp	Arg 5	Thr	Glu	Leu	Glu	Asn 10	Arg	Leu	Lys	Asp	Leu 15	Tyr		
Thr			Cys 20	Glu				25	Ile				30	Ala			
		35	Thr				40					45					
	50		Thr			55					60						,
Glu	Leu	Leu	Lys	Lys	Thr	Tyr	Glu	Thr	Ser	Leu	ser	GLU	тте	ьys	пув		

65				_	70					75					80	
Ser	His	Glu	Met	Glu 85		Lys	Ser	Leu	Glu 90	Asp	Leu	Leu	Asn	Glu 95	Lys	
Gln	Glu	Ser	Leu 100	Glu	Lys	Gln	Ile	Asn 105	Asp	Leu	Lys	Ser	Glu 110	Asn	Asp	
Ala	Leu	Asn 115	Glu	Arg	Leu											
<210 <211 <212 <213	> 37 > DN	ΙA	apie	ens												
	> CI		(1	1600))											
<400																60
cagt	gtga	tg t	ggtt	caga	ag go	caget	tcta	a gad	cctgo	ragg	aggg	gagai	eca (catto	cagagg gccagc	60 120
															ggcatc	180
tttc	ctct	ga d	ctgta	attt	ct to	ggcct	tgaa	a ga	gtaci	gag	ttta	aaaaa	aga (cagta	atgtga	240
cagt	ccat	gg a	aatt	geet	c tt	ctgt	tgaaa	a tc	tege	cacc	tgct	ccga	aag a	Me	g ttg et Leu L	298
tta	tct	ccc	aaa	ttc	tcc	tta	tcc	acc	att	cac	ata	cga	ctg	acg	gcc	346
Leu	Ser	Pro 5	Lys	Phe	Ser	Leu	Ser 10	Thr	Ile	His	Ile	Arg 15	Leu	Thr	Ala	
aaa Lys	gga Gly	ttg Leu	ctt Leu	cga Arg	aac Asn	ctt Leu	cga Arg	ctt Leu	cct Pro	tca Ser	ggg Gly	ttt Phe	agg Arg	aga Arg	agc Ser	394
	20					25					30					
act	gtt	gtt	ttc	cac	aca	gtt	gaa	aag	agc	agg	caa	aag	aat	cct -	cga	442
Thr 35	Val	Val	Phe	His	Thr 40	Val	Glu	Lys	Ser	Arg 45	Gln	Lys	Asn	Pro	Arg 50	
agc	tta	tgt	atc	cag	cca	cag	aca	gct	ccc	gat	gcg	ctg	ccc	cct	gag	490
Ser	Leu	Cys	Ile	Gln 55	Pro	Gln	Thr	Ala	Pro 60		Ala	Leu	Pro	Pro 65	Glu	
aaa	aca	ctt	gaa	ttg	acg	caa	tat	aaa	aca	aaa	tgt	gaa	aac	caa	agt	538
ГÀЗ	Thr	Leu	Glu 70	Leu	Thr	Gln	Tyr	Lys 75	Thr	Lys	Cys	Glu	Asn 80	Gln	Ser	
gga	ttt	atc	ctg	cag	ctc	aag	cag	ctt	ctt	gcc	tgt	ggt	aat	acc	aag	586
Gly	Phe	Ile 85	Leu	GIn	Leu	гув	Gln 90	ьeu	⊥eu	ΑТЯ	cys	95	ASII	Thr	пур	
ttt	gag	gca	ttg	aca	gtt	gtg	att	cag	cac	ctg	ctg	tct	gag	cgg	gag	634
Phe	Glu 100	Ala	Leu	Thr	Val	Val 105	Ile	Gln	His	Leu	Leu 110	Ser	Glu	Arg	Glu	

													gtt Val			682
													tta Leu			730
													cag Gln 160			778
													ttt Phe			826
													gca Ala			874
													gcg Ala			922
													ctt Leu			970
													aaa Lys 240			1018
													aag Lys			1066
													gat Asp			1114
													aga Arg			1162
gca Ala	aat Asn	ttg Leu	aaa Lys	aat Asn 295	cct Pro	cag Gln	atc Ile	atg Met	tat Tyr 300	cta Leu	gaa Glu	cag Gln	gag Glu	tta Leu 305	gaa Glu	1210
													cat His 320			1258

gac atc aag tta atg aaa atg gag aaa ctg gtg gac aac aac aca gca Asp IIe Lys Leu Wet Lys Met Glu Lys Leu Val Asp Asn Asn Thr Ala 325 330 330 335 ttg gtt gac aaa ttg aag cgt ttc cag cag gag aat gaa gaa ttg aaa
ttg gtt gac aaa ttg aag cgt ttc cag cag gag aat gaa gaa ttg aaa 1354 Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Glu Leu Lys 340 gct cgg atg gac aag cac atg gca atc tca agg cag ctt tcc acg gag 1402 Ala Arg Met Asp Lys His Met Ala 11e Ser Arg Gln Leu Ser Thr Glu 355 a60 a65 a770 cag gct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aca aag 1450 Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 385 cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Trp Lys Leu His Asn Gly 390 gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala 11e Pro Leu 405 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 31594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 425 aga taga cacgtccca aagtccacag actctctgaa agcattttga tgcagctcg aat cga cacgtccca aagagtac ccacagagcac ttaaagaaccy agattttaaaaa agaatttcg 1770 gaagactttt gcctccgtcc aaagagttc tccaaaaaac ctcaaaaaa agagattct gagagcttt cacacaagac gacgttgtt cacaaagacc ttaaagaaccy agattttaaaaa aagatttcg 1830 catcgaacac gacgttgtt cacaaagaact tcaagagccc aaatgagtt atcaagactt gtgaggtgaa aacgtgtgc acaaagacc ttaaagaaccy agattttaaaaa aagatttcg 1950 attaaatgtc aaaactgaa tattcagca aagagtac acagggtag aagagaggtaa atcaacaca gacgtgtgt gacacaacacca acacacaca accacacacacacaca
ttg gtt gac aaa ttg aag cgt ttc cag cag gag aat gaa gaa ttg aaa Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Glu Leu Lys 340 gct cgg atg gac aag cac atg gca atc tca agg cag ctt tcc acg gag Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser Thr Glu 355 360 agg ct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aac aag Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 375 aga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Trp Lys Leu His Asn Gly 390 agac ctg tgt agc cc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 cag tca cca agg aat tcg ggc tct tc cct agc ccc agc att tca ccc Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 aga tga cacgtccca aagtccacag acttcttgaa agcattttga tgcaggtctg Arg ** 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagacac gtgtgatcac cgagagtat tggaggtaac tggaggctgac ccaccggcgg aatcgagctt ctaaagaaca ttaaaaa aagattcg gagaggtat atcagaacatt ggcagcttgt cacacagaac ttaaagaacg agagtcttt gcctccgtcc aaaagattc tccaaaaaaa gattcaaaaa agattctg gagaggtat atcagaacttt gcctccgtcc aaaagattc tcaaggacct attaaagaacg aagatcttg gagaggttt cacacagaa acttctttga agagcttt ctaaaaa aagattcg gagaggtat atcacaca aagactttttga tgcagctt cttttccact aagactttttga tgcagcttc tttttcacct aagacttaaag ggaaaaactc tcaggacct attaaagaacg aagatcttt gttcattgc gagaggtgt atcacacac gacgtttttga cacacagaca ttaaagaacg aagacttttt gttcattgc gagaggtgt ataaaaa agagatctt gttcattgc gagaggtgt ataacacac cacctcttttg gagattttcaaaaa aagatttcg tttcattgc gagtgtgaatga aatggatgt acaagagtgt atgtgtctaa atcaataata atcaacat caccaccacac cacctctgtt aacaagaacac tcaaaaaaaa atcaatcaat cacaccacac
Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Leu Lys Lys Lis Asp Lys His Met Ala Ile Ser Arg Gln Leu Gln Leu Gln Leu Gln Leu Gln Leu Gln Lys Lu Gln Lys Gln Lys Val Asn Lys Jann J
Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Leu Lys Lys Lis Asp Lys His Met Ala Ile Ser Arg Gln Leu Gln Leu Gln Leu Gln Leu Gln Leu Gln Lys Lu Gln Lys Gln Lys Val Asn Lys Jann J
340 345 350 gct cgg atg gac aag cac atg gca atg gca atg gca gcg ctg gag lad and all all all all ser and served all all all all all all all all all al
Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser Thr Glu 370 cag gct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aac aag 1450 Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 375 cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 gac ctg ttg agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asn Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 cag tca cca agg aat tcg ggc tcc ttc cct agc cac agc att tca ccc 31594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg caggactgac cccaaggagg aacgtgggca caaagaggtat atcaggactat gtcagagggtagac tgaagacttt gcctaggagg aatcggacga ctaagagtct tcaagaacac gtgtgatcac caccggcgg aatcgacac tgaagacttt gcctcataaaaa agaatttcg cacacaag actctctgaa agcattttga tgcaggtctg 1770 gaagactttt gcctccgtc aaaagattc tccaaaaaaa gatttaaaaaa aagatttcgg 1830 catataaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat cattaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat caccaccacacaccaccaccaccaccaccaccaccacca
Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser Thr Glu 370 cag gct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aac aag 1450 Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 375 cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 gac ctg ttg agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asn Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 cag tca cca agg aat tcg ggc tcc ttc cct agc cac agc att tca ccc 31594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg caggactgac cccaaggagg aacgtgggca caaagaggtat atcaggactat gtcagagggtagac tgaagacttt gcctaggagg aatcggacga ctaagagtct tcaagaacac gtgtgatcac caccggcgg aatcgacac tgaagacttt gcctcataaaaa agaatttcg cacacaag actctctgaa agcattttga tgcaggtctg 1770 gaagactttt gcctccgtc aaaagattc tccaaaaaaa gatttaaaaaa aagatttcgg 1830 catataaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat cattaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat caccaccacacaccaccaccaccaccaccaccaccacca
355 360 365 370 cag gct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aac aag 1450 1450 Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 375 380 385 385 cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 1498 1498 Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 395 400 1546 Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 1546 Cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 1594 1650 aga tga cacgtcecca aagtccacag actctctgaa agcattttg gcaggtctg tggaggtaac tggaggtaac tggaggtaac tggaggtaac tggaggtaac tggaggtaac tggaggtaac tggaggtaac tggaggtaac tcacaagaac ttaaagaacg gattttaaaaa aggatttcgg aatcgaggt tcacaaagaac ttaaagaacg agacatttt gcttattgc 1890 1710 cattaacaca aggattt cacaaagaac ttaaagaacg acttaagaacg acttaagaacg acttaagaacg acttaagaacg acttaagaacg agacatttt taaagaact tcacacactttgc 1890 1830 cattaaatgtc aaatctga aatatcgaa atatatctga tatatctga tatgtactaa tcaaataata atatatcaat tcaaataac 220 1200 gdtgaatga aatggatgtc aaaatctga tatatctga tatgtgtctga tcaaaataata atataacactttga tcaacacacac caactctgct cagccctga aagaaaaagaa atatatcaaataa aagaattatgta aatggaggaa aacttaacaa
cag gct gtt ctg caa gag tcg ctg gag aag gag tcg aaa gtc aac aag 1450 Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val Asn Lys 375 cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg caggactgac cccaaggagg aacgtgggca caagaggtat atcagacac gtgtgtacac 1710 cgtaggtaac tggaggtca ccacggggg aatcgagct ctgaagactgg aagtctggag 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagaccac gtgtgatcac 1710 cgtaggtaac tggaggtca ccacggggg aatcgagct ctgaagactgg aagtctggag 1770 gaagactttt gcctccgcc aaagactc tcaagaacac ttaaagaacg aggattttcac agcataagg ggaaaaact tcaaggccct attaagattt atcaacctttg 1950 tttttcacct aagcataagg ggaaaaact tcagggccct attaagattt atcaacctttg 1950 tttttcacct aagaatgatc acagagtgtc acagagtgtc attagactaa tcaacatttg 1950 gtgtgaatga aatggatgtc acagagtgtc attattgtaa tcaaaataat ataacctttg 1950 attaaatgtc aaaatctgaa tattatctgga tattgtactaa tcaaaataata atcaatcaat 2130 cagcatatac attcaagcac aagccataga agaaaaagca acaggtgtgggggggggg
Capacity
Capacity
cga ctc tct atg gaa aac gag gag ctt ctg tgg aaa ctg cac aat ggg 1498 Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 395 400 gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc ctt ttg 1546 Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat cgcaggtat cgcacaggagttt gcctcgtcc aaagagttcc tccaaaaaaa gattctgaggtattte gcctcgtcc aaagagttcc tcaagaacac ttaaagaacg gacgttgtt gcacaaagacc ttaagaacg actctctgaa agcattttga tgcaggtag 1830 catcgacacg gacgttgtt cacaaagacc tcaagaggtat tcaagaacg gagttgtct cacaaagacc tcaagagatt tcaaggacct gaggtgggggggaaacact tcaaggacct tcaagagattc tcaagaacg gaggttgtca cacaagacac tcaagagattc tcaagaacg gagttttaaaaa aagattctgg 1830 catcgacata aagcataaga aacgtggtc acaagagatc tcaagagcct tcaagagcct tcaagagcct tcaagagcct tcaagagcct tcaagagcct tcaagagcct tcaagagcct tcaagagctt tcaagaacg gaggtgggggggggg
Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 395 400 gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg 1546 Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 Pro Ser Pro Ser Ile Ser Pro 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caacgaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcagacttt gcctccgtc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgtt cacaaagacac ttcaagaacg agagtattaaaaa aagatttcgg 1830 catcgacacg gacgttgtt cacaaagacac ttcaagaacg agagttttaaaaa aagatttcgg 1830 catcgacacg aacgtagag aaccttcttga aacctttgg 1950 taatgtctt cacacagac accttcttg aagactta tcaagaccac tctgagggggggggg
Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His Asn Gly 390 395 400 gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg 1546 Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 Pro Ser Pro Ser Ile Ser Pro 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caacgaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcagacttt gcctccgtc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgtt cacaaagacac ttcaagaacg agagtattaaaaa aagatttcgg 1830 catcgacacg gacgttgtt cacaaagacac ttcaagaacg agagttttaaaaa aagatttcgg 1830 catcgacacg aacgtagag aaccttcttga aacctttgg 1950 taatgtctt cacacagac accttcttg aagactta tcaagaccac tctgagggggggggg
gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg 1546 Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 700 aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgaggtct cctgagactgg aagtctgggg actgagacttt gcctccgtc aaaagattc tccaaaaaaa gagttaaaaa aagattcgg 1770 gaagacttt gcctccgtc aaaagattc tcaagaacg gagtattaaaaa aagattcgg 1830 catcgacacg gacgttgtg cacaaagcac ttaaagaacg acttttaaga aggacatct gttcattgcc 1890 tttttcacct aaggatga accttcttgg aatggtgtca tataagatt ataacctttg 1950 taatgttct caccacagac accttcttgg aagtgttcaa tcaaataata atcaatcaat 1950 attaaatgtc cacacagac accttcttgg aagtgttcaa tcaaataata tcaaataata 2130 cagcataac atttcagca aagcataga gaaaaagca atagtgtct gagaggagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtc atacattcc tgtatcttt ggggggtaact tcttggcagt 2250 gctttgactt gtccctgtc atacattcc tgtatcttt ggggggtaact tcttggcagt 2250 gctttgactt gtccctgtc atacattcc tgtatcttt ggggggtaact tctttggcagt 2310
gac ctg tgt agc ccc aag aga tcc ccc aca tcc tcc gcc atc cct ttg Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 70 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgaggtct ctgaagacttt gcctcgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagattcgg 1830 catcgacacg gacgttgtt caccacagac ttaaagaacg actctcttga aggacttct gttcattgcc 1890 ttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgtctt caccacagac accttcttgt gagtttcag tgcagcctc tctgctgtg attaaatgtc cacacagac accttcttgt gagtttcag tgcagcctc tctgctgtg attaaatgtc aaaactgaa tatatctgga tatgtctaa tcaaataata accaactaat 1310 cagcatatac attcagca aagcataga agaaaaagca acaacagggtag gagggggggggg
Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile Pro Leu 405 410 415 cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataaccttg 1950 taatgttct caccacagac acctctttt gagtttcag tctgactgg gggtggggg 2010 gtgtgaatga aatggatgc acagatgtc atgtgctga tgcagcccc tctcgtgtg 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac attcagcca caactctgct cagccctgta acagggtagg gagagggtat aacagggaga 2250 gctttgactt gtccctgct atacattctc tgtatcttt gggggtaact tcttggcagt 2310
cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc 1594 Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggegg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa agatttcgg 1830 catcgacacg gacgttgtg cacaaagcac ttaaagaacg agagatttttaacct aagcataagg ggaaaaactc tcagggcct attaagatt ataacctttg 1950 ttatgtctt caccacagac accttcttg gagtttcag tcgagctct tctgactgtg ggggtgggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tcaaataata atcaatcaat 2130 cagcatatac atttcagca aagccataga agaaaaagca acaggggtat aacaggggag 2250 gctttgactt gtccctgtct atacattctc tgtatcttt gggggtaact tcttggcagt 2310
cag tca cca agg aat tcg ggc tcc ttc cct agc ccc agc att tca ccc Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagattccg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatct gtcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaaagatt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagtttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagttc tagtgctga tgcagcctcc tctgctgtgt attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac atttcagcca aagccataga agaaaaagca gagagggtat aacaggaaga gctttgactt gtccctgtct atacattctc tgtatctttt gggggtaact tcttggcagt 2310
Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagacac ttaaagaacg agagcatctt gtcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac attcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atacattctc tgtatctttt ggggggtaact tcttggcagt 2310
Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile Ser Pro 420 425 430 aga tga cacgtcccca aagtccacag actctctgaa agcattttga tgcaggtctg 1650 Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagacac ttaaagaacg agagcatctt gtcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac attcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atacattctc tgtatctttt ggggggtaact tcttggcagt 2310
aga tga cacgtccca aagtccacag actctctgaa agcattttga tgcaggtctg Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagct ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gtcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct ataacttctc tgtatcttt gggggtaact tcttggcagt 2310
Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atcattctc tgtatcttt gggggtaact tcttggcagt 2310
Arg * 435 caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atcattctc tgtatcttt gggggtaact tcttggcagt 2310
caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gtcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtg 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct ataacttct tgtatcttt gggggtaact tcttggcagt 2310
caggactgac cccaaggagg aacgtgggca caagaggtat atcagcacac gtgtgatcac 1710 cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct ggccttgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct ataacttct tgtatcttt gggggtaact tcttggcagt 2310
cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atcactctc tgtatcttt gggggtaact tcttggcagt 2310
cgtaggtaac tggagcgtca ccaccggcgg aatcgagctt ctgagactgg aagtctggag 1770 gaagactttt gcctccgtcc aaaagattcc tccaaaaaaa gatttaaaaa aagatttcgg 1830 catcgacacg gacgttgttg cacaaagcac ttaaagaacg agagcatctt gttcattgcc 1890 tttttcacct aagcataagg ggaaaaactc tcagggccct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtgt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atcactctc tgtatcttt gggggtaact tcttggcagt 2310
categacaeg gaegttgttg cacaaagcae ttaaagaaeg agageatett gtteattgee 1890 ttttteacet aageataagg ggaaaaaete teagggeet attaagattt ataacetttg 1950 taatgttett caceacagae acettettgt gagtttteag tetgaettgtg ggggtggggg 2010 gtgtgaatga aatggatgte acagagtgte atgtgtetga tgeageetee tetgetgtt 2070 attaaatgte aaaatetgaa tatatetgga tatgtaetaa teaaataata ateaateaat 2130 cageatatae atteageea aagecataga agaaaaagea atagttgett gaattatgat 2190 catetaeeae caaetetget gteetgte ataeattete tgtatettt gggggtaaeet tettggeagt 2310
tttttcacct aagcataagg ggaaaaactc tcagggcct attaagattt ataacctttg 1950 taatgttctt caccacagac accttcttgt gagttttcag tctgactgtg ggggtggggg 2010 gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtt 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac atttcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtcctgtct atacattctc tgtatcttt gggggtaact tcttggcagt 2310
taatgttett caccacagac accttettgt gagtttteag tetgactgtg ggggtggggg 2010 gtgtgaatga aatggatgte acagagtgte atgtgtetga tgeageetee tetgetgtt 2070 attaaatgte aaaatetgaa tatatetgga tatgtactaa teaaataata ateaateaat 2130 cagcatatac attteageea aageeataga agaaaaagea atagttgett gaattatgat 2190 catetaceae caactetget cagceetgta acagggtagg gagagggtat aacaggaaga 2250 getttgaett gteeetgtet atacattete tgtatettt gggggtaact tettggeagt 2310
gtgtgaatga aatggatgtc acagagtgtc atgtgtctga tgcagcctcc tctgctgtg 2070 attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac atttcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atacattctc tgtatcttt ggggggtaact tcttggcagt 2310
attaaatgtc aaaatctgaa tatatctgga tatgtactaa tcaaataata atcaatcaat 2130 cagcatatac atttcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atacattctc tgtatctttt ggggggtaact tcttggcagt 2310
cagcatatac atttcagcca aagccataga agaaaaagca atagttgctt gaattatgat 2190 catctaccac caactctgct cagccctgta acagggtagg gagagggtat aacaggaaga 2250 gctttgactt gtccctgtct atacattctc tgtatctttt gggggtaact tcttggcagt 2310
catctaccac caactetget cageeetgta acagggtagg gagagggtat aacaggaaga 2250 getttgaett gteeetgtet atacattete tgtatettt gggggtaact tettggeagt 2310
getttgactt gteeetgtet atacattete tgtatetttt gggggtaaet tettggeagt 2310
the transfer transfer transfer transfer transfer to the transfer t
telecaging teageracy cayeryadae tagarrette tytagarret teageradee 257
atgtgagcct aacactatcc tgtaattcat tttctcaggc tatgtgtaaa tgtagaaccc 2430
taattttttt ataaaaaaac aaactaacta actgtgtaaa gaaagaaaaa gggaagtacc 2490
aacygycee ecodocoda ceesasooo gaoosasoo eyonyaaraa
thouseage thatteteat throughtha cettletee ceategagag ceacaaaage 2610
Cocceca cacceca cocceca assessed control of the con
aaacetteta ceteetaeet aettteetet gggacaagga taaaggaata tgattteea 2670
aaaccttcta cctcctacct acttttctct gggacaagga taaaggaata tgattttcca 2670 gagcccaga gccagctcat cttccaggtg ctgaaaccac tttccaaata aactaaagcc 2730 tggatttgat attacaaatt ttgggaaatc ttagaataaa gaacgagaac aaggaagtca 2790
aaaccttcta cctcctacct acttttctct gggacaagga taaaggaata tgattttcca 2670 gagcccaga gccagctcat cttccaggtg ctgaaaccac tttccaaata aactaaagcc 2730

```
gtctttggga aagtagcaga atggtccgct tctttcccat gagtggaaaa tgtggcttgt
                                                                   2970
ccaactetee tecaggitge atticagitt etitecaaaa etiattaeet eccetaatee
                                                                   3030
tgagactttg gaaaaggtgg aaggaagaac tgttgcttta tctccccctc cctgcatgtg
tcaacattgt gatgtcagta tttactaatc tacattcagt ggctgtacaa ataacagctg
                                                                   3150
tagtaagaag agattcagga tgctagaggt gaatatttgg gtcatttaca tgtacactac
                                                                   3210
atagcaagtt gatactcatg ttgcatgttc ttttaaatta gtgattttgt gtcttaagtc
                                                                   3270
tttaacttcc aatacttcat catgtatgta accttccatg tttgcttctg ataaatggaa
                                                                   3330
atgtaggttc actgccactt catgagatat ctctgctcac gcttccaagt tgttctcaat
                                                                   3390
gacattagcc aaagttgggt ttgccattca tcccctaggc atggtaaatc ttgtgttgtt
                                                                   3450
ccctgctgtc ctccgtatta cgtgaccggc aaataaatct catagcagtt aatataaaac
                                                                   3510
atctttggag gatgggagag aacaggaggg aagatgggaa acaaaataga gaattcttaa
                                                                   3570
gattttgttt aaaccaaatg tttcatgtag aatgcaaaat gttggcacgt caaaaatatg
                                                                   3630
                                                                   3690
aatgtgtaga caactgtagt tgtgctcagt ttgtagtgat gggaagtgta ttttactctg
                                                                   3742
<210> 8
<211> 435
<212> PRT
<213> Homo sapiens
<400> 8
Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Ile Arg Leu
Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Phe Arg
                               25
Arg Ser Thr Val Val Phe His Thr Val Glu Lys Ser Arg Gln Lys Asn
                           40
Pro Arg Ser Leu Cys Ile Gln Pro Gln Thr Ala Pro Asp Ala Leu Pro
                                           60
                       55
Pro Glu Lys Thr Leu Glu Leu Thr Gln Tyr Lys Thr Lys Cys Glu Asn
                                       75
                   70
Gln Ser Gly Phe Ile Leu Gln Leu Lys Gln Leu Leu Ala Cys Gly Asn
                                   90
               85
Thr Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu
                               105
Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val
                           120
Asn Leu Arg Gly Glu Leu Val Thr Ala Ser Thr Thr Cys Glu Lys Leu
                                           140
                       135
Glu Lys Ala Arg Asn Glu Leu Gln Thr Val Tyr Glu Ala Phe Val Gln
                                       155
Gln His Gln Ala Glu Lys Thr Glu Arg Glu Asn Arg Leu Lys Glu Phe
                                   170
Tyr Thr Arg Glu Tyr Glu Lys Leu Arg Asp Thr Tyr Ile Glu Glu Ala
                               185
Glu Lys Tyr Lys Met Gln Leu Gln Glu Gln Phe Asp Asn Leu Asn Ala
                                               205
                           200
        195
His Glu Thr Ser Lys Leu Glu Ile Glu Ala Ser His Ser Glu Lys Leu
                       215
Glu Leu Leu Lys Lys Ala Tyr Glu Ala Ser Leu Ser Glu Ile Lys Lys
                                       235
                   230
Gly His Glu Ile Glu Lys Lys Ser Leu Glu Asp Leu Leu Ser Glu Lys
                                   250
```

```
Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu Lys Ser Glu Asn Asp
                                265
Ala Leu Asn Glu Lys Leu Lys Ser Glu Glu Gln Lys Arg Arg Ala Arg
                            280
Glu Lys Ala Asn Leu Lys Asn Pro Gln Ile Met Tyr Leu Glu Gln Glu
                                            300
                        295
Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys Asn Glu Lys Leu His
                   310
                                        315
Gln Gln Asp Ile Lys Leu Met Lys Met Glu Lys Leu Val Asp Asn Asn
                                   330
               325
Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Glu
                                345
Leu Lys Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser
                            360
Thr Glu Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val
                        375
Asn Lys Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His
                                        395
                   390
Asn Gly Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile
                                    410
Pro Leu Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile
                                425
Ser Pro Arg
       435
<210> 9
<211> 1308
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)...(1308)
                                                                       48
atq ttq ttq tct ccc aaa ttc tcc tta tcc acc att cac ata cga ctg
Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Ile Arg Leu
acg gcc aaa gga ttg ctt cga aac ctt cga ctt cct tca ggg ttt agg
                                                                       96
Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Phe Arg
                                                                      144
aga agc act gtt gtt ttc cac aca gtt gaa aag agc agg caa aag aat
Arg Ser Thr Val Val Phe His Thr Val Glu Lys Ser Arg Gln Lys Asn
cct cga agc tta tgt atc cag cca cag aca gct ccc gat gcg ctg ccc
                                                                      192
Pro Arg Ser Leu Cys Ile Gln Pro Gln Thr Ala Pro Asp Ala Leu Pro
                         55
    50
                                                                      240
cct gag aaa aca ctt gaa ttg acg caa tat aaa aca aaa tgt gaa aac
Pro Glu Lys Thr Leu Glu Leu Thr Gln Tyr Lys Thr Lys Cys Glu Asn
```

65					70					75					80	
caa Gln	agt Ser	gga Gly	ttt Phe	atc Ile 85	ctg Leu	cag Gln	ctc Leu	aag Lys	cag Gln 90	ctt Leu	ctt Leu	gcc Ala	tgt Cys	ggt Gly 95	aat Asn	288
acc Thr	aag Lys	ttt Phe	gag Glu 100	gca Ala	ttg Leu	aca Thr	gtt Val	gtg Val 105	att Ile	cag Gln	cac His	ctg Leu	ctg Leu 110	tct Ser	gag Glu	336
														ctt Leu		384
														aaa Lys		432
gaa Glu 145	aaa Lys	gcc Ala	agg Arg	aat Asn	gag Glu 150	tta Leu	caa Gln	aca Thr	gtg Val	tat Tyr 155	gaa Glu	gca Ala	ttc Phe	gtc Val	cag Gln 160	480
														gag Glu 175		528
tac Tyr	acc Thr	agg Arg	gag Glu 180	tat Tyr	gaa Glu	aag Lys	ctt Leu	cgg Arg 185	gac Asp	act Thr	tac Tyr	att Ile	gaa Glu 190	gaa Glu	gca Ala	576
gag Glu	aag Lys	tac Tyr 195	aaa Lys	atg Met	caa Gln	ttg Leu	caa Gln 200	gag Glu	cag Gln	ttt Phe	gac Asp	aac Asn 205	tta Leu	aat Asn	gcg Ala	624
														aaa Lys		672
														aag Lys		720
ggc Gly	cat His	gaa Glu	ata Ile	gaa Glu 245	aag Lys	aaa Lys	tcg Ser	ctt Leu	gaa Glu 250	gat Asp	tta Leu	ctt. Leu	tct Ser	gag Glu 255	aag Lys	768
cag Gln	gaa Glu	tcg Ser	cta Leu 260	gag Glu	aag Lys	caa Gln	atc Ile	aat Asn 265	gat Asp	ctg Leu	aag Lys	agt Ser	gaa Glu 270	aat Asn	gat Asp	816
gct Ala	tta Leu	aat Asn 275	gaa Glu	aaa Lys	ttg Leu	aaa Lys	tca Ser 280	gaa Glu	gaa Glu	caa Gln	aaa Lys	aga Arg 285	aga Arg	gca Ala	aga Arg	864

									atc Ile							912	
									atc Ile							960	
									gag Glu 330							1008	
									ttc Phe							1056	
									gca Ala							1104	
									ctg Leu							1152	
									gag Glu							1200	
									tcc Ser 410							1248	
									tcc Ser							1296	
	ccc Pro	aga Arg 435	tga *													1308	
<210)> 10)															
<211	l> 43 2> PH	35															
			sapie	ens							١,						
)> 1(Leu		Ser	Pro	Lys	Phe	Ser	Leu	Ser	Thr	Ile	His	Ile	Arg	Leu		
1				5					10 Arg					15			
			20					25					30				
arg	ser	Inr	vaı	٧aı	Fue	HIS	III	Vdl	Glu	пλа	Set	ALG	GIII	пур	VPII		

```
40
Pro Arg Ser Leu Cys Ile Gln Pro Gln Thr Ala Pro Asp Ala Leu Pro
Pro Glu Lys Thr Leu Glu Leu Thr Gln Tyr Lys Thr Lys Cys Glu Asn
                  70
Gln Ser Gly Phe Ile Leu Gln Leu Lys Gln Leu Leu Ala Cys Gly Asn
                                   90
Thr Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu
Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val
                           120
Asn Leu Arg Gly Glu Leu Val Thr Ala Ser Thr Thr Cys Glu Lys Leu
                       135
Glu Lys Ala Arg Asn Glu Leu Gln Thr Val Tyr Glu Ala Phe Val Gln
                                       155
                  150
Gln His Gln Ala Glu Lys Thr Glu Arg Glu Asn Arg Leu Lys Glu Phe
                                   170
               165
Tyr Thr Arq Glu Tyr Glu Lys Leu Arg Asp Thr Tyr Ile Glu Glu Ala
                               185
Glu Lys Tyr Lys Met Gln Leu Gln Glu Gln Phe Asp Asn Leu Asn Ala
                           200
His Glu Thr Ser Lys Leu Glu Ile Glu Ala Ser His Ser Glu Lys Leu
                                           220
                       215
Glu Leu Leu Lys Lys Ala Tyr Glu Ala Ser Leu Ser Glu Ile Lys Lys
                                       235
                   230
Gly His Glu Ile Glu Lys Lys Ser Leu Glu Asp Leu Leu Ser Glu Lys
                                   250
               245
Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu Lys Ser Glu Asn Asp
                               265
Ala Leu Asn Glu Lys Leu Lys Ser Glu Glu Gln Lys Arg Arg Ala Arg
                           280
Glu Lys Ala Asn Leu Lys Asn Pro Gln Ile Met Tyr Leu Glu Gln Glu
                       295
Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys Asn Glu Lys Leu His
                                       315
                  310
Gln Gln Asp Ile Lys Leu Met Lys Met Glu Lys Leu Val Asp Asn Asn
                                   330
               325
Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Glu
                              345
Leu Lys Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser
                           360
Thr Glu Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val
                       375
Asn Lys Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His
                                       395
                   390
Asn Gly Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile
                                   410
               405
Pro Leu Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile
                               425
Ser Pro Arg
       435
```

<210> 11

<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Oligonucleotide primer	
<400>	11	
cgcgga	tccc agacagaccg gacggaactg gag	33
<210>	12	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Oligonucleotide primer	
<400>	12	
ccqqaa	uttca ctacaacctt tcqtttaaaq catc	34

Rec'd PCT/PTO 19 APR 2001

1

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
- (A) NAME: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE-CNRS
 - (B) STREET: 3 rue Michel-Ange
 - (C):CITY: Paris
 - (E) COUNTRY: FRANCE
 - (F) POSTAL CODE: 75794 PARIS Cedex 16
 - (A) NAME: ELBAZ Nathalie
 - (B) STREET: 7 Passage des Italiens
 - (C) CITY: Bagnolet
 - (E) COUNTRY: FRANCE
 - (F) POSTAL CODE: 93170
 - (A) NAME: NAHMIAS Clara
 - (B) STREET: 4 rue Bailly
 - (C) CITY: Paris
 - (E) COUNTRY: FRANCE
 - (F) POSTAL CODE: 75003
 - (A) NAME: STROSBERG Arthur Donny
 - (B) STREET: 66 rue de Javel
 - (C) CITY: Paris
 - (E) COUNTRY: FRANCE
 - (F) POSTAL CODE: 75015
- (ii) TITLE OF THE INVENTION: NUCLEIC SEQUENCES ENCODING AN AT2 RECEPTOR-INTERACTING PROTEIN (ATIP) AND THEIR APPLICATIONS
 - (iii) NUMBER OF SEQUENCES: 12
 - (iv) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
- (2) INFORMATION FOR SEQ ID NO: 1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1803 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION:178..1500

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCTACCCCC CCCCA	ACGCAC CCCCCAATC	T GGGTGGCCTG	GCATTAGCAT GTAAGCTTG	r 60
TTTTCTCTGG CTGTA	ATCTCT TGGCCTGGA	A GAACCCCGAG	TTGCCAAGAG ACACAGTATO	G 120
TGATGGTCCC TGGAA	AAAGCT GCTTCCCCT	G CGAAGTTCTC	CCACTGGCTT CGAAGAC	177
			ATC CAC GTC CGC CTA Ile His Val Arg Leu 15	225
			CCT TCG GGG CTC AGG Pro Ser Gly Leu Arg 30	273
			GGC AGG CAG AAG AAT Gly Arg Gin Lys Asn 45	321
			CCA GAT GTG CTG TCC Pro Asp Val Leu Ser 60	369
			ACA AAA TGT GAA AGC Thr Lys Cys Glu Ser 80	417
			CTT TCC CGT GGT AAC Leu Ser Arg Gly Asn 95	465
			CAC CTC CTG TCT GAG His Leu Leu Ser Glu 110	513
			TCT CAA GAA CTT GTC Ser Gln Glu Leu Val 125	561
			GCC TGT GAG AAG CTA Ala Cys Glu Lys Leu 140	609
			CAA GAA TTT GTC CAG Gln Glu Phe Val Gln 160	657
Lys Leu Asn Gln			GAA CTG GAG AAC CGG Glu Leu Glu Asn Arg 175	705
			CTT CAG AGC ATT TAC Leu Gln Ser Ile Tyr 190	753

ATT Ile	GAG Glu	GAG Glu 195	GCA Ala	GAA Glu	AAA Lys	TAT Tyr	AAA Lys 200	ACT Thr	CAA Gln	CTG Leu	CAA Gln	GAG Glu 205	CAG Gln	TTT Phe	GAC Asp	801
AAC Asn	TTA Leu 210	AAC Asn	GCC Ala	GCC Ala	CAT His	GAG Glu 215	ACC Thr	ACT Thr	AAG Lys	Leu	GAG Glu 220	ATT Ile	GAA Glu	GCT Ala	AGC Ser	849
CAC His 225	TCG Ser	GAG Glu	AAG Lys	GTG Val	GAA Glu 230	TTG Leu	CTG Leu	AAG Lys	AAG Lys	ACC Thr 235	TAT Tyr	GAA Glu	ACC Thr	TCC Ser	CTT Leu 240	897
TCA Ser	GAA Glu	ATC Ile	AAG Lys	AAG Lys 245	AGC Ser	CAT His	GAG Glu	ATG Met	GAG Glu 250	AAG Lys	AAG Lys	TCA Ser	CTG Leu	GAG Glu 255	GAT Asp	945
CTG Leu	CTT Leu	AAT Asn	GAG Glu 260	AAG Lys	CAG Gln	GAA Glu	TCG Ser	CTG Leu 265	GAG Glu	AAA Lys	CAA Gln	ATC Ile	AAT Asn 270	GAT Asp	CTG Leu	993
AAG Lys	AGT Ser	GAA Glu 275	AAC Asn	GAT Asp	GCT Ala	TTA Leu	AAC Asn 280	GAA Glu	AGG Arg	TTG Leu	AAA Lys	TCA Ser 285	GAG Glu	GAG Glu	CAA Gln	1041
AAG Lys	CAA Gln 290	CTG Leu	TCA Ser	AGA Arg	GAG Glu	AAG Lys 295	GCG Ala	AAT Asn	TCC Ser	AAA Lys	AAC Asn 300	CCT Pro	CAG Gln	GTC Val	ATG Met	1089
TAT Tyr 305	CTG Leu	GAG Glu	CAA Gln	GAA Glu	CTA Leu 310	GAA Glu	AGC Ser	CTG Leu	AAG Lys	GCT Ala 315	GTG Val	TTA Leu	GAG Glu	ATC Ile	AAG Lys 320	1137
AAT Asn	GAG Glu	AAG Lys	CTG Leu	CAC His 325	CAG Gln	CAG Gln	GAC Asp	ATG Met	AAG Lys 330	CTA Leu	ATG Met	AAG Lys	ATG Met	GAA Glu 335	AAG Lys	1185
CTG Leu	GTG Val	GAC Asp	AAT Asn 340	AAC Asn	ACA Thr	GCA Ala	TTG Leu	GTT Val 345	GAC Asp	AAG Lys	CTG Leu	AAG Lys	CGA Arg 350	TTC Phe	CAG Gln	1233
CAG Gln	Glu	AAC Asn 355	Glu	Glu	Leu	Lys	Ala	CGC Arg	Met	Asp	Lys	His	Met	GCA Ala	ATT Ile	1281
TCA Ser	AGG Arg 370	Gln	CTT Leu	TCC Ser	ACC Thr	GAG Glu 375	CAG Gln	GCC Ala	GCG Ala	CTG Leu	CAA Gln 380	GAG Glu	TCC Ser	CTT Leu	GAG Glu	1329
AAG Lys 385	Glu	TCA Ser	AAG Lys	GTC Val	AAC Asn 390	Lys	AGA Arg	CTG Leu	TCC	ATG Met 395	Glu	AAC Asn	GAG Glu	GAA Glu	CTT Leu 400	1377
CTG Leu	TGG Trp	AAA Lys	CTG Leu	CAC His 405	Asn	GGA Gly	GAC Asp	CTG Leu	TGC Cys 410	Ser	CCC Pro	AAG Lys	AGA Arg	TCC Ser 415	CCC Pro	1425

ACC TCC TCG GCC ATC CCT TTC CAG TCC CCC AGG AAT TCT GGT TCC TTC Thr Ser Ser Ala Ile Pro Phe Gln Ser Pro Arg Asn Ser Gly Ser Phe 420 425 430	1473
TCC AGC CCC AGC ATC TCA CCC AGA TGA CGGCTTCTGA ACGCAGGAGA Ser Ser Pro Ser Ile Ser Pro Arg * 435 440	1520
CTCTCTGAAG GCACTGAGGT GCGCTTCTGC AGGACTGACC CTCTCATGGG AACTCGAGTT	1580
GCTGCGTTAG CTCTCTGGAA TATCCCCAGG ATATCGGGAG AGCAGCCGCC AACCGTATCA	1640
GCTACGTACG AATAGAGAGC TCCAATAGAA GACTTTTAAC TTGGTCCAAA AGCCTCCTCC	1700
AAAAACAGAT TTCGGAACTG AAGTGGACAT AGTTGCACAA AGCACTTACG GAACGAGGGA	1760
ACCTTGTTCT TTGCCTTCCT TCACCTAAGC ATAGGCTTTC CAG	1803

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 440 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
- Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Val Arg Leu
 1 5 10 15
- Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Leu Arg 20 25 30
- Lys Asn Thr Val Ile Phe His Thr Val Glu Lys Gly Arg Gln Lys Asn 35 40 45
- Pro Arg Ser Leu Cys Ile Gln Thr Gln Thr Ala Pro Asp Val Leu Ser 50 55 60
- Ser Glu Arg Thr Leu Glu Leu Ala Gln Tyr Lys Thr Lys Cys Glu Ser 65 70 75 80
- Gln Ser Gly Phe Ile Leu His Leu Arg Gln Leu Leu Ser Arg Gly Asn 85 90 95
- Asn Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu
 100 105 110
- Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val 115 120 125
- Ser Leu Arg Gly Glu Leu Val Ala Ala Ser Ser Ala Cys Glu Lys Leu 130 135 140

Glu 145	Lys	Ala	Arg	Ala	Asp 150	Leu	Gln	Thr	Ala	Tyr 155	Gln	Glu	Phe	Val	Gln 160
Lys	Leu	Asn	Gln	Gln 165	His	Gln	Thr	Asp	Arg 170	Thr	Glu	Leu	Glu	Asn 175	Arg
Leu	Lys	Asp	Leu 180	Tyr	Thr	Ala	Glu	Cys 185	Glu	Lys	Leu	Gln	Ser 190	Ile	Tyr
Ile	Glu	Glu 195	Ala	Glu	Lys	Tyr	Lys 200	Thr	Gln	Leu	Gln	Glu 205	Gln	Phe	Asp
Asn	Leu 210	Asn	Ala	Ala	His	Glu 215	Thr	Thr	Lys	Leu	Glu 220	Ile	Glu	Ala	Ser
His 225	Ser	Glu	Lys	Val	Glu 230	Leu	Leu	Lys	Lys	Thr 235	Tyr	Glu	Thr	Ser	Leu 240
Ser	Glu	Ile	Lys	Lys 245	Ser	His	Glu	Met	Glu 250	Lys	Lys	Ser	Leu	Glu 255	Asp
Leu	Leu	Asn	Glu 260	Lys	Gln	Glu	Ser	Leu 265	Glu	Lys	Gln	Ile	Asn 270	Asp	Leu
Lys	Ser	Glu 275	Asn	Asp	Ala	Leu	Asn 280	Glu	Arg	Leu	Lys	Ser 285	Glu	Glu	Gln
Lys	Gln 290	Leu	Ser	Arg	Glu	Lys 295	Ala	Asn	Ser	Lys	Asn 300	Pro	Gln	Val	Met
Tyr 305	Leu	Glu	Gln	Glu	Leu 310	Glu	Ser	Leu	Lys	Ala 315	Val	Leu	Glu	Ile	Lys 320
Asn	Glu	Lys	Leu	His 325	Gln	Gln	Asp	Met	Lys 330	Leu	Met	Lys	Met	Glu 335	Lys
Leu	Val	Asp	Asn 340	Asn	Thr	Ala	Leu	Val 345	Asp	Lys	Leu	Lys	Arg 350	Phe	Gln
Gln	Glu	Asn 355	Glu	Glu	Leu	Lys	Ala 360	Arg	Met	Asp	Lys	His 365	Met	Ala	Ile
Ser	Arg 370	Gln	Leu	Ser	Thr	Glu 375	Gln	Ala	Ala	Leu	Gln 380	Glu	Ser	Leu	Glu
Lys 385	Glu	Ser	Lys	Val	Asn 390	Lys	Arg	Leu	Ser	Met 395	Glu	Asn	Glu	Glu	Leu 400
Leu	Trp	Lys	Leu	His 405	Asn	Gly	Asp	Leu	Cys 410	Ser	Pro	Lys	Arg	Ser 415	Pro
Thr	Ser	Ser	Ala 420	Ile	Pro	Phe	Gln	Ser 425	Pro	Arg	Asn	Ser	Gly 430	Ser	Phe
Ser	Ser	Pro 435	Ser	Ile	Ser	Pro	Arg 440	*							

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1323 base pairs (B) TYPE: nucleotide

 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION:1..1322
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATG Met	CTG Leu	TTG Leu	TCT Ser 445	CCC Pro	AAA Lys	TTC Phe	TCC Ser	TTA Leu 450	TCC Ser	ACC Thr	ATC Ile	CAC His	GTC Val 455	CGC Arg	CTA Leu	48
ACC Thr	GCC Ala	AAA Lys 460	GGA Gly	CTG Leu	CTT Leu	CGA Arg	AAC Asn 465	CTC Leu	CGG Arg	CTT Leu	CCT Pro	TCG Ser 470	GGG Gly	CTC Leu	AGG Arg	96
AAA Lys	AAC Asn	ACT Thr	GTC Val	ATT Ile	TTC Phe	CAC His	ACA Thr	GTT Val	GAA Glu	AAG Lys	GGC Gly	AGG Arg	CAG Gln	AAG Lys	AAT Asn	144
CCC Pro	AGG Arg	AGC Ser	CTG Leu	TGC Cys	ATC Ile	CAG Gln	ACC Thr	CAG Gln	ACA Thr	GCT Ala	CCA Pro	GAT Asp	GTG Val	CTG Leu	TCC Ser	192
TCC Ser	GAG Glu	AGA Arg	ACG Thr	CTT Leu	GAG Glu	TTG Leu	GCC Ala	CAA Gln	TAC Tyr	AAG Lys	ACA Thr	AAA Lys	TGT Cys	GAA Glu	AGC Ser	240
CAA Gln	AGT Ser	GGA Gly	TTC Phe	ATC Ile	CTG Leu	CAC His	CTC Leu	AGG Arg	CAG Gln	CTT Leu	CTT Leu	TCC	CGT Arg	GGT Gly	AAC Asn	288
AAC Asn	AAG Lys	TTT Phe	GAA Glu	GCG Ala	CTG Leu	ACA Thr	GTT Val	GTG Val	ATC Ile	CAG Gln	CAC His	CTG Leu	CTG Leu	TCT Ser	GAG Glu	336
CGG Arg	GAG Glu	GAA Glu	GCA Ala	CTG Leu	AAG Lys	CAA Gln	CAC His	AAA Lys	ACC Thr	CTC Leu	TCT Ser	CAA Gln	GAA Glu	CTT Leu	GTC Val	384
AGC Ser	CTC Leu	CGG Arg	GGA Gly	GAG Glu	CTA Leu	GTT Val	GCT Ala	GCT Ala	TCA Ser	AGC Ser	GCC Ala	TGT Cys	GAG Glu	AAG Lys	CTA Leu	432
GAA Glu	AAG Lys	GCT Ala	AGG Arg	GCT Ala	GAC Asp	TTA Leu	CAG Gln	ACA Thr	GCG Ala	TAT Tyr	CAA Gln	GAA Glu	TTT Phe	GTC Val	CAG Gln	480

AAA	CTA	AAC	CAG	CAG	CAT	CAG	ACA	GAC	CGG	ACG	GAA	CTG	GAG	AAC	CGG	528
Lys	Leu	Asn	Gln	Gln	His	Gln	Thr	Asp	Arg	Thr	Glu	Leu	Glu	Asn	Arg	
CTG	AAG	GAC	TTA	TAC	ACC	GCA	GAG	TGT	GAG	AAG	CTT	CAG	AGC	ATT	TAC	576
Leu	Lys	Asp	Leu	Tyr	Thr	Ala	Glu	Cys	Glu	Lys	Leu	Gln	Ser	Ile	Tyr	
ATT	GAG	GAG	GCA	GAA	AAA	TAT	AAA	ACT	CAA	CTG	CAA	GAG	CAG	TTT	GAC	624
Ile	Glu	Glu	Ala	Glu	Lys	Tyr	Lys	Thr	Gln	Leu	Gln	Glu	Gln	Phe	Asp	
AAC	TTA	AAC	GCC	GCC	CAT	GAG	ACC	ACT	AAG	CTT	GAG	ATT	GAA	GCT	AGC	672
Asn	Leu	Asn	Ala	Ala	His	Glu	Thr	Thr	Lys	Leu	Glu	Ile	Glu	Ala	Ser	
CAC	TCG	GAG	AAG	GTG	GAA	TTG	CTG	AAG	AAG	ACC	TAT	GAA	ACC	TCC	CTT	720
His	Ser	Glu	Lys	Val	Glu	Leu	Leu	Lys	Lys	Thr	Tyr	Glu	Thr	Ser	Leu	
TCA	GAA	ATC	AAG	AAG	AGC	CAT	GAG	ATG	GAG	AAG	AAG	TCA	CTG	GAG	GAT	768
Ser	Glu	Ile	Lys	Lys	Ser	His	Glu	Met	Glu	Lys	Lys	Ser	Leu	Glu	Asp	
CTG	CTT	AAT	GAG	AAG	CAG	GAA	TCG	CTG	GAG	AAA	CAA	ATC	AAT	GAT	CTG	816
Leu	Leu	Asn	Glu	Lys	Gln	Glu	Ser	Leu	Glu	Lys	Gln	Ile	Asn	Asp	Leu	
AAG	AGT	GAA	AAC	GAT	GCT	TTA	AAC	GAA	AGG	TTG	AAA	TCA	GAG	GAG	CAA	864
Lys	Ser	Glu	Asn	Asp	Ala	Leu	Asn	Glu	Arg	Leu	Lys	Ser	Glu	Glu	Gln	
AAG	CAA	CTG	TCA	AGA	GAG	AAG	GCG	AAT	TCC	AAA	AAC	CCT	CAG	GTC	ATG	912
Lys	Gln	Leu	Ser	Arg	Glu	Lys	Ala	Asn	Ser	Lys	Asn	Pro	Gln	Val	Met	
TAT	CTG	GAG	CAA	GAA	CTA	GAA	AGC	CTG	AAG	GCT	GTG	TTA	GAG	ATC	AAG	960
Tyr	Leu	Glu	Gln	Glu	Leu	Glu	Ser	Leu	Lys	Ala	Val	Leu	Glu	Ile	Lys	
AAT	GAG	AAG	CTG	CAC	CAG	CAG	GAC	ATG	AAG	CTA	ATG	AAG	ATG	GAA	AAG	1008
Asn	Glu	Lys	Leu	His	Gln	Gln	Asp	Met	Lys	Leu	Met	Lys	Met	Glu	Lys	
- CTG	GTG	GAC	AAT	AAC	ACA	GCA	TTG	GTT	GAC	AAG	CTG	AAG	CGA	TTC	CAG	1056
Leu	Val	Asp	Asn	Asn	Thr	Ala	Leu	Val	Asp	Lys	Leu	Lys	Arg	Phe	Gln	
Gln	Glu	Asn	Glu	Glu	TTA Leu	Lys	Ala	Arg	Met	Asp	Lys	His	Met	Ala	Ile	1104
TCA	AGG	CAA	CTT	TCC	ACC	GAG	CAG	GCC	GCG	CTG	CAA	GAG	TCC	CTT	GAG	1152
Ser	Arg	Gln	Leu	Ser	Thr	Glu	Gln	Ala	Ala	Leu	Gln	Glu	Ser	Leu	Glu	

								GAA Glu		1200
								TCC Ser		1248
								TCC Ser		1296
 		ATC Ile		TG	A					1323

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 440 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Val Arg Leu
1 5 10 15

Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Leu Arg 20 25 30

Lys Asn Thr Val Ile Phe His Thr Val Glu Lys Gly Arg Gln Lys Asn \$35\$ 40 45

Pro Arg Ser Leu Cys Ile Gln Thr Gln Thr Ala Pro Asp Val Leu Ser 50 55 60

Ser Glu Arg Thr Leu Glu Leu Ala Gln Tyr Lys Thr Lys Cys Glu Ser 65 70 75 80

Gln Ser Gly Phe Ile Leu His Leu Arg Gln Leu Leu Ser Arg Gly Asn 85 90 95

Asn Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu
100 105 110

Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val 115 120 125

Ser Leu Arg Gly Glu Leu Val Ala Ala Ser Ser Ala Cys Glu Lys Leu 130 135 140

Glu Lys Ala Arg Ala Asp Leu Gln Thr Ala Tyr Gln Glu Phe Val Gln 145 150 155 160 Lys Leu Asn Gln Gln His Gln Thr Asp Arg Thr Glu Leu Glu Asn Arg 165 170 175

Leu Lys Asp Leu Tyr Thr Ala Glu Cys Glu Lys Leu Gln Ser Ile Tyr 180 185 190

Ile Glu Glu Ala Glu Lys Tyr Lys Thr Gln Leu Gln Glu Gln Phe Asp 195 200 205

Asn Leu Asn Ala Ala His Glu Thr Thr Lys Leu Glu Ile Glu Ala Ser 210 215 220

His Ser Glu Lys Val Glu Leu Leu Lys Lys Thr Tyr Glu Thr Ser Leu 225 230 235 240

Ser Glu Ile Lys Lys Ser His Glu Met Glu Lys Lys Ser Leu Glu Asp 245 250 255

Leu Leu Asn Glu Lys Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu 260 265 270

Lys Ser Glu Asn Asp Ala Leu Asn Glu Arg Leu Lys Ser Glu Glu Gln 275 280 285

Lys Gln Leu Ser Arg Glu Lys Ala Asn Ser Lys Asn Pro Gln Val Met 290 295 300

Tyr Leu Glu Gln Glu Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys 305 310 315 320

Asn Glu Lys Leu His Gln Gln Asp Met Lys Leu Met Lys Met Glu Lys 325 330 335

Leu Val Asp Asn Asn Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln 340 345 350

Gln Glu Asn Glu Glu Leu Lys Ala Arg Met Asp Lys His Met Ala Ile 355 360 365

Ser Arg Gln Leu Ser Thr Glu Gln Ala Ala Leu Gln Glu Ser Leu Glu 370 375 380

Lys Glu Ser Lys Val Asn Lys Arg Leu Ser Met Glu Asn Glu Glu Leu 385 390 395 400

Leu Trp Lys Leu His Asn Gly Asp Leu Cys Ser Pro Lys Arg Ser Pro 405 410 415

Thr Ser Ser Ala Ile Pro Phe Gln Ser Pro Arg Asn Ser Gly Ser Phe 420 425 430

Ser Ser Pro Ser Ile Ser Pro Arg 435 440

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 354 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..354
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CAT His	CAG Gln	ACA Thr	GAC Asp 440	CGG Arg	ACG Thr	GAA Glu	CTG Leu	GAG Glu 445	AAC Asn	CGG Arg	CTG Leu	AAG Lys	GAC Asp 450	TTA Leu	TAC Tyr	48
ACC Thr	GCA Ala	GAG Glu 455	TGT Cys	GAG Glu	AAG Lys	CTT Leu	CAG Gln 460	AGC Ser	ATT Ile	TAC Tyr	ATT Ile	GAG Glu 465	GAG Glu	GCA Ala	GAA Glu	96
AAA Lys	TAT Tyr 470	AAA Lys	ACT Thr	CAA Gln	CTG Leu	CAA Gln 475	GAG Glu	CAG Gln	TTT Phe	GAC Asp	AAC Asn 480	TTA Leu	AAC Asn	GCC Ala	GCC Ala	144
CAT His 485	GAG Glu	ACC Thr	ACT Thr	AAG Lys	CTT Leu 490	GAG Glu	ATT Ile	GAA Glu	GCT Ala	AGC Ser 495	CAC His	TCG Ser	GAG Glu	AAG Lys	GTG Val 500	192
GAA Glu	TTG Leu	CTG Leu	AAG Lys	AAG Lys 505	ACC Thr	TAT Tyr	GAA Glu	ACC Thr	TCC Ser 510	CTT Leu	TCA Ser	GAA Glu	ATC Ile	AAG Lys 515	AAG Lys	240
AGC Ser	CAT His	GAG Glu	ATG Met 520	GAG Glu	AAG Lys	AAG Lys	TCA Ser	CTG Leu 525	GAG Glu	GAT Asp	CTG Leu	CTT Leu	AAT Asn 530	GAG Glu	AAG Lys	288
CAG Gln	GAA Glu	TCG Ser 535	CTG Leu	GAG Glu	AAA Lys	CAA Gln	ATC Ile 540	AAT Asn	GAT Asp	CTG Leu	AAG Lys	AGT Ser 545	GAA Glu	AAC Asn	GAT Asp	336
GCT	TTA	AAC	GAA	AGG	TTG											354

(2) INFORMATION FOR SEQ ID NO: 6:

Ala Leu Asn Glu Arg Leu

550

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 118 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

	(ii)	MOI	LECUI	LE TY	PE:	prot	ein										
	(xi)	SEC	QUEN	CE DE	ESCRI	PTIC	ON: 5	SEQ :	D NO): 6:	;						
His 1	Gln	Thr	Asp	Arg 5	Thr	Glu	Leu	Glu	Asn 10	Arg	Leu	Lys	Asp	Leu 15	Tyr		
Thr	Ala	Glu	Cys 20	Glu	Lys	Leu	Gln	Ser 25	Ile	Tyr	Ile	Glu	Glu 30	Ala	Glu		
Lys	Tyr	Lys 35	Thr	Gln	Leu	Gln	Glu 40	Gln	Phe	Asp	Asn	Leu 45	Asn	Ala	Ala		
His	Glu 50	Thr	Thr	Lys	Leu	Glu 55	Ile	Glu	Ala	Ser	His 60	Ser	Glu	Lys	Val		
Glu 65	Leu	Leu	Lys	Lys	Thr 70	Tyr	Glu	Thr	Ser	Leu 75	Ser	Glu	Ile	Lys	Lys 80		
Ser	His	Glu	Met	Glu 85	Lys	Lys	Ser	Leu	Glu 90	Asp	Leu	Leu	Asn	Glu 95	Lys		
Gln	Glu	Ser	Leu 100	Glu	Lys	Gln	Ile	Asn 105	Asp	Leu	Lys	Ser	Glu 110	Asn	Asp		
Ala	Leu	Asn 115	Glu	Arg	Leu												
(2)	INF	AMAC	rion	FOR	SEQ	ID 1	10: °	7:									
	(i	() () ()	A) L: B) T' C) S'	CE CI ENGTI YPE: TRANI OPOLO	H: 3' nucl	742 l Leot: ESS:	oase ide sing	pai:	rs								
	(ii) MO	LECU	LE T	YPE:	CDN	A										
	(ix	(E: AME/I OCAT:			.160	0									
	(xi) SE	QUEN	CE D	ESCR:	[PTI	ON:	SEQ	ID N	D: 7	:						
CAG	TGTG.	ATG '	TGGT	TCAG	AG G	CAGC'	TTCT.	A GA	CCTG	CAGG	AGG	gaga'	TTG '	TATT	CAGAG	G .	6
AAG	AGCA	TCA	TTTT	GGCA	AC A	rctg	AAAGʻ	T GA	AAAC	GGAA	GCC	AGAA	ACA	CTTG	GCCAG	3C	12
CCT	GGGG	GAT	TTTT	TTCT'	TC T	ATGC	CTCT	G TG	GTGG	AATG	ACA	TTTG	CTG	TGTA	GGCAT	.c	18
TTT	CCTC	TGA	CTGT	ATTT	CT TO	GCC'	TTGA.	A GA	GTAC'	TGAG	TTT.	AAAA	AGA	CAGT	ATGTG	S A	24
CAG	TCCA	TGG .	TAAA	TGCC	TC T	rctg'	TGAA.	A TC	TCGC	CACC	TGC	rccg.	AAG .		TG et		29

															-	
TTG Leu	TTG Leu	TCT Ser	CCC Pro	AAA Lys	TTC Phe	TCC Ser	TTA Leu	TCC Ser	ACC Thr	ATT Ile	CAC His	ATA Ile	CGA Arg	CTG Leu	ACG Thr	343
GCC Ala	AAA Lys	GGA Gly	TTG Leu	CTT Leu	CGA Arg	AAC Asn	CTT Leu	CGA Arg	CTT Leu	CCT Pro	TCA Ser	GGG Gly	TTT Phe	AGG Arg	AGA Arg	391
AGC Ser	ACT Thr	GTT Val	GTT Val	TTC Phe	CAC His	ACA Thr	GTT Val	GAA Glu	AAG Lys	AGC Ser	AGG Arg	CAA Gln	AAG Lys	AAT Asn	CCT Pro	439
CGA Arg	AGC Ser	TTA Leu	TGT Cys	ATC Ile	CAG Gln	CCA Pro	CAG Gln	ACA Thr	GCT Ala	CCC Pro	GAT Asp	GCG Ala	CTG Leu	CCC	CCT Pro	487
GAG Glu	AAA Lys	ACA Thr	CTT Leu	GAA Glu	TTG Leu	ACG Thr	CAA Gln	TAT Tyr	AAA Lys	ACA Thr	AAA Lys	TGT Cys	GAA Glu	AAC Asn	CAA Gln	535
AGT Ser	GGA Gly	TTT Phe	ATC Ile	CTG Leu	CAG Gln	CTC Leu	AAG Lys	CAG Gln	CTT Leu	CTT Leu	GCC Ala	TGT Cys	GGT Gly	AAT Asn	ACC Thr	583
AAG Lys	TTT Phe	GAG Glu	GCA Ala	TTG Leu	ACA Thr	GTT Val	GTG Val	ATT Ile	CAG Gln	CAC His	CTG Leu	CTG Leu	TCT Ser	GAG Glu	CGG Arg	631
GAG Glu	GAA Glu	GCA Ala	CTG Leu	AAA Lys	CAA Gln	CAC His	AAA Lys	ACC Thr	CTA Leu	TCT Ser	CAA Gln	GAA Glu	CTT Leu	GTT Val	AAC Asn	679
CTC Leu	CGG Arg	GGA Gly	GAG Glu	CTA Leu	GTC Val	ACT Thr	GCT Ala	TCA Ser	ACC Thr	ACC Thr	TGT Cys	GAG Glu	AAA Lys	TTA Leu	GAA Glu	727
AAA Lys	GCC Ala	AGG Arg	AAT Asn	GAG Glu	TTA Leu	CAA Gln	ACA Thr	GTG Val	TAT Tyr	GAA Glu	GCA Ala	TTC Phe	GTC Val	CAG Gln	CAG Gln	775
CAC His	CAG Gln	GCT Ala	GAA Glu	AAA Lys	ACA Thr	GAA Glu	CGA Arg	GAG Glu	AAT Asn	CGG Arg	CTT Leu	AAA Lys	GAG Glu	TTT Phe	TAC Tyr	823
ACC Thr	AGG Arg	GAG Glu	TAT Tyr	GAA Glu	AAG Lys	CTT Leu	CGG Arg	GAC Asp	ACT Thr	TAC Tyr	ATT Ile	GAA Glu	GAA Glu	GCA Ala	GAG Glu	871
AAG Lys	TAC Tyr	AAA Lys	ATG Met	CAA Gln	TTG Leu	CAA Gln	GAG Glu	CAG Gln	TTT Phe	GAC Asp	AAC Asn	TTA Leu	AAT Asn	GCG Ala	CAT His	919
GAA Glu	ACC Thr	TCT Ser	AAG Lys	TTG Leu	GAA Glu	ATT Ile	GAA Glu	GCT Ala	AGC Ser	CAC His	TCA Ser	GAG Glu	AAA Lys	CTT Leu	GAA Glu	967
TTG Leu	CTA Leu	AAG Lys	AAG Lys	GCC Ala	TAT Tyr	GAA Glu	GCC Ala	TCC Ser	CTT Leu	TCA Ser	GAA Glu	ATT Ile	AAG Lys	AAA Lys	GGC Gly	1015

CAT	GAA	ATA	GAA	AAG	AAA	TCG	CTT	GAA	GAT	TTA	CTT	TCT	GAG	AAG	CAG	1063
His	Glu	Ile	Glu	Lys	Lys	Ser	Leu	Glu	Asp	Leu	Leu	Ser	Glu	Lys	Gln	
GAA	TCG	CTA	GAG	AAG	CAA	ATC	AAT	GAT	CTG	AAG	AGT	GAA	AAT	GAT	GCT	1111
Glu	Ser	Leu	Glu	Lys	Gln	Ile	Asn	Asp	Leu	Lys	Ser	Glu	Asn	Asp	Ala	
TTA	AAT	GAA	AAA	TTG	AAA	TCA	GAA	GAA	CAA	AAA	AGA	AGA	GCA	AGA	GAA	1159
Leu	Asn	Glu	Lys	Leu	Lys	Ser	Glu	Glu	Gln	Lys	Arg	Arg	Ala	Arg	Glu	
AAA	GCA	AAT	TTG	AAA	AAT	CCT	CAG	ATC	ATG	TAT	CTA	GAA	CAG	GAG	TTA	1207
Lys	Ala	Asn	Leu	Lys	Asn	Pro	Gln	Ile	Met	Tyr	Leu	Glu	Gln	Glu	Leu	
GAA	AGC	CTG	AAA	GCT	GTG	TTA	GAG	ATC	AAG	AAT	GAG	AAA	CTG	CAT	CAA	1255
Glu	Ser	Leu	Lys	Ala	Val	Leu	Glu	Ile	Lys	Asn	Glu	Lys	Leu	His	Gln	
CAG	GAC	ATC	AAG	TTA	ATG	AAA	ATG	GAG	AAA	CTG	GTG	GAC	AAC	AAC	ACA	1303
Gln	Asp	Ile	Lys	Leu	Met	Lys	Met	Glu	Lys	Leu	Val	Asp	Asn	Asn	Thr	
GCA Ala	TTG Leu	GTT Val	GAC Asp	AAA Lys	TTG Leu	AAG Lys	CGT	TTC Phe	CAG Gln	CAG Gln	GAG Glu	AAT Asn	GAA Glu	GAA Glu	TTG Leu	1351
AAA	GCT	CGG	ATG	GAC	AAG	CAC	ATG	GCA	ATC	TCA	AGG	CAG	CTT	TCC	ACG	1399
Lys	Ala	Arg	Met	Asp	Lys	His	Met	Ala	Ile	Ser	Arg	Gln	Leu	Ser	Thr	
GAG	CAG	GCT	GTT	CTG	CAA	GAG	TCG	CTG	GAG	AAG	GAG	TCG	AAA	GTC	AAC	1447
Glu	Gln	Ala	Val	Leu	Gln	Glu	Ser	Leu	Glu	Lys	Glu	Ser	Lys	Val	Asn	
AAG	CGA	CTC	TCT	ATG	GAA	AAC	GAG	GAG	CTT	CTG	TGG	AAA	CTG	CAC	AAT	1495
Lys	Arg	Leu	Ser	Met	Glu	Asn	Glu	Glu	Leu	Leu	Trp	Lys	Leu	His	Asn	
GGG Gly	GAC Asp	CTG Leu	TGT Cys	AGC Ser	CCC Pro	AAG Lys	AGA Arg	TCC	CCC Pro	ACA Thr	TCC Ser	TCC Ser	GCC Ala	ATC Ile	CCT Pro	1543
TTG	CAG	TCA	CCA	AGG	AAT	TCG	GGC	TCC	TTC	CCT	AGC	CCC	AGC	ATT	TCA	1591
Leu	Gln	Ser	Pro	Arg	Asn	Ser	Gly	Ser	Phe	Pro	Ser	Pro	Ser	Ile	Ser	
	AGA Arg		CAC	GTCC	CCA .	AAGT	CCAC.	AG A	CTCT	CTGA	A AG	CATT	TTGA			1640
TGC	AGGT	CTG	CAGG	ACTG	AC C	CCAA	GGAG	g aa	CGTG	GGCA	CAA	GAGG	TAT	ATCA	GCACAC	1700
GTG	TGAT	CAC	CGTA	GGTA	AC T	GGAG	CGTC	A CC	ACCG	GCGG	AAT	CGAG	CTT	CTGA	GACTGG	1760

					-	
AAGTCTGGAG	GAAGACTTTT	GCCTCCGTCC	AAAAGATTCC	TCCAAAAAA	GATTTAAAAA	1820
AAGATTTCGG	CATCGACACG	GACGTTGTTG	CACAAAGCAC	TTAAAGAACG	AGAGCATCTT	1880
GTTCATTGCC	TTTTTCACCT	AAGCATAAGG	GGAAAAACTC	TCAGGGCCCT	ATTAAGATTT	1940
ATAACCTTTG	TAATGTTCTT	CACCACAGAC	ACCTTCTTGT	GAGTTTTCAG	TCTGACTGTG	2000
GGGGTGGGGG	GTGTGAATGA	AATGGATGTC	ACAGAGTGTC	ATGTGTCTGA	TGCAGCCTCC	2060
TCTGCTGTGT	ATTAAATGTC	AAAATCTGAA	TATATCTGGA	TATGTACTAA	TCAAATAATA	2120
ATCAATCAAT	CAGCATATAC	ATTTCAGCCA	AAGCCATAGA	AGAAAAAGCA	ATAGTTGCTT	2180
GAATTATGAT	CATCTACCAC	CAACTCTGCT	CAGCCCTGTA	ACAGGGTAGG	GAGAGGGTAT	2240
AACAGGAAGA	GCTTTGACTT	GTCCCTGTCT	ATACATTCTC	TGTATCTTTT	GGGGGTAACT	2300
TCTTGGCAGT	TTTTCAGTGT	TCAGCCATGT	CAGTTGAAAC	TAGATTTTTC	TGTAGATTTT	2360
TTACTTACCC	ATGTGAGCCT	AACACTATCC	TGTAATTCAT	TTTCTCAGGC	TATGTGTAAA	2420
TGTAGAACCC	TAATTTTTCT	АТАААААААС	AAACTAACTA	ACTGTGTAAA	GAAAGAAAAA	2480
GGGAAGTACC	AATGGGTTTT	TCCACCTTAT	TTTTACCTTT	GATCTACCCT	TGCAGATTTA	2540
ACCTGTCTTC	TTCCCTCCCA	TTATTCTCAT	TTTCCTTTTA	CCTTTCTCCA	CCATCCAGAG	2600
CCACAAAAGC	AAACCTTCTA	CCTCCTACCT	ACTTTTCTCT	GGGACAAGGA	TAAAGGAATA	2660
TGATTTTCCA	GAGCCCCAGA	GCCAGCTCAT	CTTCCAGGTG	CTGAAACCAC	TTTCCAAATA	2720
AACTAAAGCC	TGGATTTGAT	ATTACAAATT	TTGGGAAATC	TTAGAATAAA	GAACGAGAAC	2780
AAGGAAGTCA	TTGGCTAGTA	TAATTAAGAA	AGGTAGGATT	CAGTGCTTAC	CGATGATGCA	2840
GTACTTGATA	GAAGAAAACA	GTCTGGGAGG	ATAGCGCTCA	TTTTTCAGTT	ACCCTTTAAG	2900
GAGTCCCTTT	GTCTTTGGGA	AAGTAGCAGA	ATGGTCCGCT	TCTTTCCCAT	GAGTGGAAAA	2960
TGTGGCTTGT	CCAACTCTCC	TCCAGGTTGC	ATTTCAGTTT	CTTTCCAAAA	CTTATTACCT	3020
CCCCTAATCC	TGAGACTTTG	GAAAAGGTGG	AAGGAAGAAC	TGTTGCTTTA	TCTCCCCCTC	3080
CCTGCATGTG	TCAACATTGT	GATGTCAGTA	TTTACTAATC	TACATTCAGT	GGCTGTACAA	31,40
ATAACAGCTG	TAGTAAGAAG	AGATTCAGGA	TGCTAGAGGT	GAATATTTGG	GTCATTTACA	3200
TGTACACTAC	ATAGCAAGTT	GATACTCATG	TTGCATGTTC	TTTTAAATTA	GTGATTTTGT	3260
GTCTTAAGTC	TTTAACTTCC	AATACTTCAT	CATGTATGTA	ACCTTCCATG	TTTGCTTCTG	3320
ATAAATGGAA	ATGTAGGTTC	ACTGCCACTT	CATGAGATAT	CTCTGCTCAC	GCTTCCAAGT	3380
TGTTCTCAAT	GACATTAGCC	AAAGTTGGGT	TTGCCATTCA	TCCCCTAGGC	ATGGTAAATC	3440
TTGTGTTGTT	CCCTGCTGTC	CTCCGTATTA	CGTGACCGGC	AAATAAATCT	CATAGCAGTT	3500

AATATAAAAC	ATCTTTGGAG	GATGGGAGAG	AACAGGAGGG	AAGATGGGAA	ACAAAATAGA	3560
GAATTCTTAA	GATTTTGTTT	AAACCAAATG	TTTCATGTAG	AATGCAAAAT	GTTGGCACGT	3620
CAAAAATATG	AATGTGTAGA	CAACTGTAGT	TGTGCTCAGT	TTGTAGTGAT	GGGAAGTGTA	3680
TTTTACTCTG	ATCAAATAAA	TAATGCTGGA	ATACTCAAAA	ААААААААА	AAAAAAAAA	3740
AA						3742

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 435 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Leu Leu Ser Pro Lys Phe Ser Leu Ser Thr Ile His Ile Arg Leu 1 5 10 15

Thr Ala Lys Gly Leu Leu Arg Asn Leu Arg Leu Pro Ser Gly Phe Arg 20 25 30

Arg Ser Thr Val Val Phe His Thr Val Glu Lys Ser Arg Gln Lys Asn 35 40 45

Pro Arg Ser Leu Cys Ile Gln Pro Gln Thr Ala Pro Asp Ala Leu Pro 50 55 60

Pro Glu Lys Thr Leu Glu Leu Thr Gln Tyr Lys Thr Lys Cys Glu Asn 65 70 75 80

Gln Ser Gly Phe Ile Leu Gln Leu Lys Gln Leu Leu Ala Cys Gly Asn 85 90 95

Thr Lys Phe Glu Ala Leu Thr Val Val Ile Gln His Leu Leu Ser Glu 100 105 110

Arg Glu Glu Ala Leu Lys Gln His Lys Thr Leu Ser Gln Glu Leu Val

Asn Leu Arg Gly Glu Leu Val Thr Ala Ser Thr Thr Cys Glu Lys Leu 130 135 140

Glu Lys Ala Arg Asn Glu Leu Gln Thr Val Tyr Glu Ala Phe Val Gln 145 150 155 160

Gln His Gln Ala Glu Lys Thr Glu Arg Glu Asn Arg Leu Lys Glu Phe 165 170 175

Tyr Thr Arg Glu Tyr Glu Lys Leu Arg Asp Thr Tyr Ile Glu Glu Ala 180 185 190 Glu Lys Tyr Lys Met Gln Leu Gln Glu Gln Phe Asp Asn Leu Asn Ala 195 200 205

His Glu Thr Ser Lys Leu Glu Ile Glu Ala Ser His Ser Glu Lys Leu 210 215 220

Glu Leu Leu Lys Lys Ala Tyr Glu Ala Ser Leu Ser Glu Ile Lys Lys 225 230 235 240

Gly His Glu Ile Glu Lys Lys Ser Leu Glu Asp Leu Leu Ser Glu Lys 245 250 255

Gln Glu Ser Leu Glu Lys Gln Ile Asn Asp Leu Lys Ser Glu Asn Asp 260 265 270

Ala Leu Asn Glu Lys Leu Lys Ser Glu Glu Gln Lys Arg Arg Ala Arg 275 280 285

Glu Lys Ala Asn Leu Lys Asn Pro Gln Ile Met Tyr Leu Glu Gln Glu 290 295 300 .

Leu Glu Ser Leu Lys Ala Val Leu Glu Ile Lys Asn Glu Lys Leu His 305 . 310 315 320

Gln Gln Asp Ile Lys Leu Met Lys Met Glu Lys Leu Val Asp Asn Asn 325 330 335

Thr Ala Leu Val Asp Lys Leu Lys Arg Phe Gln Gln Glu Asn Glu Glu 340 345 350

Leu Lys Ala Arg Met Asp Lys His Met Ala Ile Ser Arg Gln Leu Ser 355 360 365

Thr Glu Gln Ala Val Leu Gln Glu Ser Leu Glu Lys Glu Ser Lys Val 370 375 380

Asn Lys Arg Leu Ser Met Glu Asn Glu Glu Leu Leu Trp Lys Leu His 385 390 395 400

Asn Gly Asp Leu Cys Ser Pro Lys Arg Ser Pro Thr Ser Ser Ala Ile 405 410 415

Pro Leu Gln Ser Pro Arg Asn Ser Gly Ser Phe Pro Ser Pro Ser Ile 420 425 430

Ser Pro Arg * 435

3

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1308 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single

1308

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9: ATGTTGTTGT CTCCCAAATT CTCCTTATCC ACCATTCACA TACGACTGAC GGCCAAAGGA 60 TTGCTTCGAA ACCTTCGACT TCCTTCAGGG TTTAGGAGAA GCACTGTTGT TTTCCACACA 120 GTTGAAAAGA GCAGGCAAAA GAATCCTCGA AGCTTATGTA TCCAGCCACA GACAGCTCCC 180 GATGCGCTGC CCCCTGAGAA AACACTTGAA TTGACGCAAT ATAAAACAAA ATGTGAAAAC 240 CAAAGTGGAT TTATCCTGCA GCTCAAGCAG CTTCTTGCCT GTGGTAATAC CAAGTTTGAG 300 GCATTGACAG TTGTGATTCA GCACCTGCTG TCTGAGCGGG AGGAAGCACT GAAACAACAC 360 AAAACCCTAT CTCAAGAACT TGTTAACCTC CGGGGAGAGC TAGTCACTGC TTCAACCACC 420 TGTGAGAAAT TAGAAAAAGC CAGGAATGAG TTACAAACAG TGTATGAAGC ATTCGTCCAG 480 CAGCACCAGG CTGAAAAAAC AGAACGAGAG AATCGGCTTA AAGAGTTTTA CACCAGGGAG 540 TATGAAAAGC TTCGGGACAC TTACATTGAA GAAGCAGAGA AGTACAAAAT GCAATTGCAA 600 GAGCAGTTTG ACAACTTAAA TGCGCATGAA ACCTCTAAGT TGGAAATTGA AGCTAGCCAC 660 TCAGAGAAAC TTGAATTGCT AAAGAAGGCC TATGAAGCCT CCCTTTCAGA AATTAAGAAA 720 GGCCATGAAA TAGAAAAGAA ATCGCTTGAA GATTTACTTT CTGAGAAGCA GGAATCGCTA 780 GAGAAGCAAA TCAATGATCT GAAGAGTGAA AATGATGCTT TAAATGAAAA ATTGAAATCA 840 900 GAAGAACAAA AAAGAAGAGC AAGAGAAAAA GCAAATTTGA AAAATCCTCA GATCATGTAT CTAGAACAGG AGTTAGAAAG CCTGAAAGCT GTGTTAGAGA TCAAGAATGA GAAACTGCAT 960 CAACAGGACA TCAAGTTAAT GAAAATGGAG AAACTGGTGG ACAACAACAC AGCATTGGTT 1020 GACAAATTGA AGCGTTTCCA GCAGGAGAAT GAAGAATTGA AAGCTCGGAT GGACAAGCAC 1080 ATGGCAATCT CAAGGCAGCT TTCCACGGAG CAGGCTGTTC TGCAAGAGTC GCTGGAGAAG 1140 GAGTCGAAAG TCAACAAGCG ACTCTCTATG GAAAACGAGG AGCTTCTGTG GAAACTGCAC 1200 AATGGGGACC TGTGTAGCCC CAAGAGATCC CCCACATCCT CCGCCATCCC TTTGCAGTCA 1260

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs

CCAAGGAATT CGGGCTCCTT CCCTAGCCCC AGCATTTCAC CCAGATGA

- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	
CAAGCGTTCT CTCGGAGGAC A	21
(2) INFORMATION FOR SEQ ID NO: 11:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleotide (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:	33
(2) INFORMATION FORSEQ ID NO: 12:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleotide (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:	
CCGGAATTCA CTACAACCTT TCGTTTAAAG CATC	34